UDK 811.134.3'373.611
https://doi.org/10.22210/suvlin.2017.083.03
Izvorni znanstveni članak
Prihvaćeno za tisak: 20. travnja 2017.

Alexandra Soares Rodrigues
ESE – Polytechnic Institute of Bragança, CELGA–ILTEC – University of Coimbra
afsr@ipb.pt

Limits on the extension of affixal combination: structural restrictions and processing conditions[1]

The study of the mental lexicon has been fostered by the analysis of the way complex words are mentally represented and processed. This paper concerns the syntagmatic extension of multiple affixation; specifically, the processing of complex words that contain four suffixes that operate in word–formation patterns of Portuguese. Although the individual addition of suffixes obeys structural constraints, the multiple combination results in complex words with low frequency and low expectedness by the speaker, which contribute to the lack of semantic transparency and of affixal salience of the combination. Our study demonstrates a relation between these factors and the experience of the speaker with the affixal combination, which determines the pattern character of the combination. We suggest that a suffix exerts the prediction of other suffixes as long as the combination is expected. Non–frequent heterocategorial complex words with a combination of four suffixes are contrasted with non–frequent words containing pleonastic affixation. In the latter type of words, the redundancy of semantic structures increases the semantic transparency of the word, which suggests a prediction effect operating on the semantic level of the affixal combination. Processing of complex words is dependent on the level of expectedness of the speaker towards the affix combination, which constrains the level of word acceptance by speakers.

1. Introduction

This paper aims to analyse the syntagmatic extension of multiple affixation (at least three suffixes) in contemporary Portuguese productive patterns (Rio–Torto et al. 2016) from the perspective of the interaction between processing and semantic and morphological structures. Following Jackendoff (2002: 34), we assume that a combination of competence and processing theories is required for a more complete understanding of language and its

1 The author would like to express her sincere gratitude to the anonymous reviewers for their useful, constructive comments on this work. The author would also like to thank Sérgio Carvalho for having checked English language. Obviously, the author is the only responsible for any remnant mistakes and incongruences.

phenomena. This theoretical principle is supported by our analysis of the syntagmatic extension of multiple affixation, whose structural limits are shown to be dependent on processing conditions. We base our findings on the analysis of corpora (Reference Corpus of Contemporary Portuguese and Linguateca), Google searches and experiments (acceptability judgment task and recall tasks) with European Portuguese native speakers.

From the perspective of the description of structural constraints, multiple affixation ought to be possible as long as structural constraints between affixes are preserved. However, we want to understand how far this multiple combination maintains semanticality, that is, semantic well–formedness (Pustejovsky 1995: 40), and grammaticality, and can be processed by native speakers.

Structural constraints predict that words found in Google such as *comercializabilidade* 'tradability, quality of being able to be commercialized' and *materializabilidade* 'quality of being able to be materialized' are possible forms. As will be shown in section 4, these are the result of multiple affixation and each one of the involved affixations operates in existent derivational patterns of contemporary Portuguese:

$matéria_{N}$>$materi_{N}$–al_{ADJ}>$material_{ADJ}$–iz_{V}>$materializ(a)_{V}$–vel_{ADJ}>$materializabil_{ADJ}$–$idade_{N}$

Despite the structural well–formedness of the derivatives that result from these patterns, according to Experiment 1 (an acceptability judgment task) forms such as *materializabilidade* are not considered acceptable by Portuguese native speakers. Results from Experiment 2, a recall task, manifest the difficulty native speakers have in processing those type of words containing a chain of four derivational suffixes that operate in different word–formation patterns.

A parallel dual–route model (Baayen, Dijkstra & Schreuder 1997; Schreuder & Baayen 1997) could account for the processing of these kinds of words, where the decomposed route would be favoured owing to the combination of two factors: the very low frequency of the word and its morphological complexity (Niswander–Klement & Pollatsek 2006). The decomposed route implies a decomposition of the stimulus into morphological constituents followed by an integration or recombination of those constituents that leads to the recognition of the whole word (Taft 2004). As demonstrated in Niswander–Klement and Pollatsker (2006), word length has a bias towards the decomposed route in a dual–route model.

In *–bil*/*–al–izabilidade* words, however, the affixal chain is much too extensive for a computational processing. Considering morphemes as either lexical items or as spell–outs of rules (cf. Beard 1995), we follow authors such as Schreuder and Baayen (1995: 132), who defend the notion that morphological structure serves the computation of meaning, Ji et al. (2011: 419), who emphasize the importance of semantics to the processing of words, and Gagné and Spalding (2004, 2009, 2010), who suggest that the reconstruction of meaning departs from information of the meanings of constituents, if this kind of information is available (cf. also Libben 1998: 32 for compounds).

The goal of morphological analysis made by the speaker is to create meaning (Libben 2015) and the presence of morphemes in a word expresses a conceptual category in word–formation. Thus, in *–bil/–al–izabilidade* words, the quantity of morphological constituents with no referential semantics contributes to the low semantic transparency and the low affixal salience (Laudanna and Burani 1995) of this affixal combination, which makes morphological parsing difficult.

There is also an obstacle for whole–word processing: because this combination of suffixes has a very low frequency, it does not correspond to predictable/expected combinations in the speaker's mind (Hawkins and Blakeslee 2004; Plag and Baayen 2009), and thus, it is not represented as a mental word–formation pattern. What is important is not the mental representation of a specific word, but the derivational pattern that permits the construction and analysis of that word.

One particular type of multiple affixation contrasts with those cases:

Affix combinations that produce evaluative nouns/adjectives such as *cas–inh–oto–zinho* 'very small and simple hut', *pequen–in–inho–zinho* 'very, very very small' are easily processed via the decomposed route. The hypotheses are as follow: 1) Evaluative affixes are characterized by a higher semantic transparency (Bell and Schäfer 2016) than suffixes such as *–al–*, *–bil–*. 2) In such evaluative formations, each suffix repeats the semantic information of its precedent. In the case of pleonastic affixation (Gardani 2015; Dressler et al. 2015; Szymanek 2015), multiple affixation is reinforcing the information that is needed to process the word.

In the contrasting situation, expectedness (Bell and Schäfer 2013; 2016) was revealed to be a crucial factor for the recognition of the complex word. Expectedness may be either towards a word–formation pattern, the combination of suffixes or semantic structures expressed by morphemes.

Variables such as frequency of combination, semantic transparency, affixal salience and expectedness are important for the pattern character of the affix combination; that is, for the capacity of a specific affix combination, as a whole, to function as a mental pattern that serves the analysis and, if productive, the construction of words. Pattern character of an affixal chain depends on the experience of the speaker and, intrinsically, on the expectedness towards the affix combination. It can be observed in two ways: i) the more a phenomenon is experienced, the stronger the mental pattern gets; and ii) the stronger the mental pattern, the easier the recognition of particular instances of the phenomenon.

1.1 Some remarks on our objective

Our work focuses on affixation and not on compounding, although many authors emphasize compounds as more interesting in psycholinguistic studies than affixed words (cf. Ji et. al. 2011: 407; Libben 2005: 270). Compounds may be considered a more challenging phenomenon for the hypothesis of decomposition into constituents. This is because: i) compounds have lower frequency

than affixed words; ii) unlike affixed words, compounds have open–class units as constituents.

Without arguing these aspects, we consider that affixation and its limits are also challenging aspects for psycholinguistics and linguistics. Although also ruled by psycholinguistic conditions, the syntagmatic limits to compounding seem to be more flexible than the limits on the number of affixes in affixation (cf. Booij (2016) on Dutch composition, Barz (2016) on German composition, Kumar et al. (2010) on Sanskrit composition and Bauer et al. (2013:507–508) and Bauer (1983: 69) for English derivational affixation).

Second, the low frequency of the types under study may be objected to in our work. There are two main reasons that provide a rationale for the study of these words: a) although they have low frequency, a number of words with the affixal chain under study do exist, which makes them language phenomena deserving of study; and b) their low frequency suggests there must be factors underpinning that phenomenon. A complete understanding of language cannot only account for frequent phenomena, and the mechanisms that underlie them. It is also important to explain why some phenomena are rare, and why some possibilities do not occur. There is no science based on the analysis of a single particular object, or of a few particular objects. This is true for astronomy, physics and biology, and should also be the case for linguistics and psycholinguistics. For example, octopuses have blue instead of red blood, an unusual phenomenon in animal species. This gives scientists two options: a) as this phenomenon is not well represented in the kingdom *animalia*, it is considered worthless to study it; or b) it is considered worthy of study to determine the reasons that underlie its rarity.

A third remark concerns the kind of mental representation that our study targets. Are we targeting the representation of complex words in the mental lexicon or the way access to those words takes place? Marslen–Wilson et al. (1994: 4) take "[...] the lexical entry to be the modality–independent core representation of a word's syntactic and semantic attributes as well as its abstract morphological properties [...]" and define the access representation of the word, as "[...] from the modality–specific access representation, [...] provides the perceptual target for lexical access, defining the route whereby information in the sensory input is linked to a given lexical entry". Our study focuses on the access of complex words interwoven with their representation in the mental lexicon. Access to a word in the mental lexicon is dependent on the way (and if) it is represented there, as may be understood by perspectives on expectedness of matches between stimuli and data in the mind (Hawkins and Blakeslee 2004; Plag and Baayen 2009).

1.2 Questions on the objective

In this section we summarise the main questions on the objective to be developed on the article and its organisation. This paper aims to analyse the syntagmatic extension of multiple affixation involving four suffixes that work on contemporary Portuguese productive patterns (Rio–Torto et al. 2016). In

Portuguese, words constructed by means of multiple affixation such as *comercializabilidade*[2] 'tradability, quality of being able to be commercialized', *contabilizabilidade*[3] 'quality of being able to be accounted', and *socializabilidade*[4] 'quality of being able to be socialized' are found in Google searches.

Our empirical consideration of these words evoked two primary questions:

a) Are the particular derivational operations involved in the formation of these particular words frequent in Portuguese?
b) Are these words structurally well–formed – that is, do the different affixation procedures obey structural constraints of Portuguese word–formation (section 4)?

A search in corpora and on Google suggests that the answer to question a) is negative (Appendix A). An analysis of the structural constraints of Portuguese and comparison of the words with them reveals the answer to the second question is affirmative (section 4). Knowing that *–bil/–al–izabilidade* words are not frequent, but do obey structural constraints, leads to further questions:

c) Although a few *–bil/–al–izabilidade* words exist, do Portuguese native speakers accept them (Experiment 1)?
d) Why are *–bil/–al–izabilidade* words not frequent and, assuming that speakers do not accept them, what are the factors that underlie the non–acceptance (Experiment 2)?

Our first hypothesis is thus: Processing conditions motivated by multiple affixation must underlie the low frequency of these words. At the same time, the low frequency of the words motivates the computation of the words to be accessed. However, there are too many affixes to be processed for a computation processing.

Following these questions and hypotheses, a new question arises:

e) Do other complex words manifest the same degree of frequency and acceptance/non–acceptance (Experiment 2)?

To answer this question, we decided to make a comparison between *–bil/–al–izabilidade* words with words presenting multiple pleonastic suffixation, specifically, evaluative suffixation.

A central question emerges:

f) Does the combination *–bil/–al–iz(a)–bil–idade* constitute a word–formation pattern? Although the words that contain this combination are

---

2 Google search (accessed on 14/04/2016): *bibliotecadigital.fgv.br/dspace/bitstream/10438/1040/1/2202.pdf* : Em muitos setores o grau de "*comercializabilidade*" (tradability) é insuficiente, devido aos altos custos de transporte em relação ao peso, à perecibilidade [...]» ('In many sectors the degree of "tradability" (tradability [sic]) is insufficient, owing to the high costs of transportation in relation with weight, to perishability [...]').

3 Google search (accessed on 14/04/2016): *https://www.linkedin.com/in/gilberto–plettes–223683104* : análises de *contabilizabilidade*, assistência a obra, solução de interferências em campo) ('analysis of "quality of being able to be accounted", assistance to work, solutions of interferences in field').

4 Google search (accessed on 14/04/2016): http://patriciapiedadesouza.blogspot.pt/2010_04_01_archive.html: ou que criam seus cães de maneira aleatória e depois vendem ou doam os filhotes sem nenhuma imunidade ou *socializabilidade*. ('or [who] breed their dogs in a random way and who then sell or give the puppies away without any immunity or quality of being able to be socialized').

not frequent but do obey structural constraints, can we describe the multiple combination of the affixes involved as a derivational pattern of Portuguese?

From this arises a final question:

g) What are the consequences for complex word processing, in the case where the combination of affixes does not constitute a word–formation pattern?

Bearing in mind the set of questions and hypotheses formulated, we briefly present theories on affix combination (section 2) and complex word processing (section 3). More specific processing factors involved in the limits of syntagmatic extension of affixation, such as semantic transparency, affixal salience, expectedness of the affixal combination and the pattern character of the combination, are discussed in section 3.1. In section 4, we describe concepts related to structural constraints and their relation to word–formation patterns of Portuguese. We analyse the correspondence of the specific affixal combination *–bil/–al–iz(a)–bil–idade* with word–formation patterns in 4.1. In section 5, we describe the acceptability judgment task and recall task, the experiments that are the empirical grounds of our study. Section 6 is dedicated to the general discussion, and in section 7 we present the main conclusions of the study.

2. Affix combination

In linguistics, the way complex words are constructed has been the focus of morphology, either of lexicalist morphology (e.g. Aronoff 1976; Halle 1973) or of other theories that situate morphology in the field of syntax (e.g. distributed morphology [cf. Halle and Marantz 1993]). We restrict our focus and theoretical background to the principles of lexicalist morphology.

One of the many concerns in the study of morphologically complex words has been the combination of affixes within the same word. In recent years, studies have focused on affix combination (cf. Hay and Plag 2004; Plag and Baayen 2009; Talamo 2015; Manova 2010b; 2011; Manova and Aronoff 2010a; Manova and Aronoff 2010b). Traditionally, combination of affixes has been explained by stratum–oriented models (Siegel 1974; Allen 1978; Kiparsky 1982; Giegerich 1999), by selectional restrictions of each particular affix (Fabb 1988; Plag 1999; Talamo 2010; Rodrigues 2015), by the interaction of selectional restrictions and processing constraints, by what is called the Complexity–Based Ordering hypothesis (formulated by Hay [2002] and applied, with different results, by Hay [2003]; Plag [2002]; Hay and Plag [2004]; Gaeta [2005]; Plag and Baayen [2009]; Manova [2010a]; Zirkel [2010]; Talamo [2010]; Saarinen and Hay [2014]), and by the interaction between scope, phonological subcategorization and morphotactic constraints (Caballero 2010), under the view of Optimality Theory (cf. also Ryan (2010). The evaluation of the balance between universal and language–specific factors that determine affix ordering has been the matter of debate for studies such as Sims and Parker (2015), Caballero (2010), Ryan (2010), among others.

Stratum–oriented models analyse combinations of affixes as determined by the lexical strata (*viz.*, Latinate *vs.* Germanic strata for languages such as English) each particular affix belongs to. Selectional restrictions, *i.e.*, structural constraints, restrain the combinations between affixes operating in word formation rules at the level of phonological, syntactic and semantic features of the affixes. Structural constraints work at the level of the grammaticality of combinations between affixes. The Complexity–Based Ordering hypothesis suggests that there is a relation between affix ordering and morphological parsing, with the affixes being closer to the derivational base more difficult to parse than those that appear in a more external position. This relation makes affixal combination predictable. Taking the approach of Optimality Theory, Caballero (2010) analyses if affix ordering is language–specific or dependent on cross–linguistic principles by proposing that affix ordering arises from an interaction between scope, phonological subcategorization and morphotactic constraints. She proposes that semantically non–compositional sequences of affixes may emerge from priming effects and morphophonologically conditioned multiple exponence. Sims and Parker (2015), following the Complexity–Based Ordering hypothesis, propose that the way lexical processing is achieved in each particular language determines the degree of affix ordering freeness. The authors demonstrate that decomposition leads to more freedom of affix ordering than whole–word processing. They show that combinability in Russian is freer than in English, because in Russian words are generally more likely to be parsed than processed as a whole during lexical processing than English words.

Neither of the approaches is dedicated to the analysis of the syntagmatic extension of affixation; that is, they are not dedicated to understanding the conditions that limit the number of derivational suffixes combined in a single word. In our study, we aim to analyse words that contain a combination of four suffixes and to contribute to a better understanding of the processing conditions involved in the possible extension of those combinations.

3. Processing of morphological complex words

One of the most studied aspects of psycholinguistics is the way the mental lexicon is represented and processed. The way complex words are represented and processed constitutes an important topic (cf. Schreuder and Baayen 1995; Baayen 2007). A longstanding debate (cf. Libben 1998: 30–31) concerns access and the mental representation of complex words. Models that advocate that complex words are accessed via morphological decomposition (computation) are contrasted with those that suggest they are lexically stored and accessed as whole units (storage) (cf. Baayen (2007), Bertram, Schreuder and Baayen (2000) and Kuperman et al. (2010) for the balance between the effects of storage vs. computation in lexical processing).

The first type, the so–called sublexical model, can be represented by studies such as Taft (1979; 1981), Andrews (1986), Andrews et al. (2004), Libben (1998), and Marslen–Wilson et al. (1994), who propose that complex words are

decomposed into their morphologic constituents so that they can be accessed. In this sense, morphological constituents are lexically stored (Taft 1981; Marslen–Wilson et al. 1994), and the access to the constituent representations can facilitate access to the whole–word representation (Zwitserlood 1994; Schreuder and Baayen 1995; 1997; Libben 1998, Taft 2003; Taft and Kougious 2004). The second model type, the so–called supralexical models, is illustrated by Butterworth (1983) and Giraudo and Grainger (2001). They propose that the normal processing of complex words does not entail decomposition into morphological constituents but direct access because these words are stored as units in the mental lexicon. Both these theories are single–route models.

A comparison of the pros and cons of each type of access has shown that a decomposed mechanism would lead to an overburden of the computational system, while storage of complex words as whole–units would lead to an over capacity in storage (cf. Baayen (2007); Kuperman et al. (2009); Kuperman, et al. (2010)). Although a direct route is most efficient for the processing of high–frequency complex words (e.g. Lehtonen et al. 2007), morphological parsing represents a gain in terms of storage (Marslen–Wilson et al. 1994). This conflict lies at the centre of theories that argue for a dual–route model (Frauenfelder and Schreuder (1991); Schreuder and Baayen (1995); Baayen, Dijkstra and Schreuder (1997)). A dual–route model allows for the possibility of complex words to be accessed via a direct route and decomposition. In the dual route model proposed by Schreuder and Baayen (1995), competition occurs between the two processes. The faster route 'wins' and gets to operate the processing of each particular word. Dual–route parallel models (such as Schreuder and Baayen (1995), Baayen and Schreuder (1999; 2000)) and models that combine morpho–orthographic and morphosemantic systems that may lead to either the whole–word or decomposition (proposed by Diependaele et al. (2005)) respond to scientific evidence that many factors influence access to complex words (cf. Kuperman et al. (2010) for a review of the area).

Studies addressing the processing of complex words as a non–linear task have attempted to account for the numerous complex factors involved. Laudanna and Burani (1995) emphasize that many factors may be involved in the way each complex word is processed. As pointed out by Libben (1998: 31), some of those factors involve "frequency, lexical category [...], morphological type (derived, inflected, compounded), and the semantic relationships between the multimorphemic forms and their constituents". Root and base frequency (Taft and Forster 1976; Beauvillain 1996; Burani and Caramazza 1987; Schreuder, Burani and Baayen 2003; Niswander, Pollatsek and Rayner 2000; Niswander–Klement and Pollatsek 2006), surface frequency (Baayen et al. 2007), semantic transparency (Marslen–Wilson et al. 1994; Libben et al. 2003), productivity (which rules the recognition of non–frequent complex words constructed by means of productive rules) (Baayen 1992; 1993), family size and family frequency of the base (Bertram, Schreuder and Baayen 2000) and affixal salience (Laudanna and Burani 1995) are examples of the factors analysed in the processing of complex words. To combine these multiple factors, models such as Marslen–Wilson et al. (1994) and Schreuder and Baayen (1995) were devel-

oped. According to Schreuder and Baayen (1995: 146), "conceptual complexity, semantic transparency, phonological transparency, the complexity of the word formation operations, pseudo–affixation, and affixal homonymy" should be taken into account in an adequate model of morphological processing.

One of the key factors shown to be involved in the decomposition of complex words is semantic transparency. Complex words with an opaque semantic structure are not processed via the decomposition of their morphological structure, whilst words that are semantically transparent are. Recently, Ji et al. (2011: 422) demonstrated that "the availability of constituent representations and the semantic integration of those representations play an important role in ease of processing". According to these authors, in compounds, morphological parsing does not take more time than direct processing when access to the constituents of the compound occurs and if the constituents are more frequent than the compound as a whole. There may be conflict in semantic integration owing to the semantics of the constituents and the stored semantic representation of the compound. This conflict retards the processing of compounds. For the decomposed route to succeed there must be conditions that enable the identification of constituents. Semantic transparency is one of those conditions.

Concerning affixed words, Laudanna and Burani (1995) proposed a conceptualization of affixal perceptual salience, "the likelihood that a derivational affix will serve as a processing unit" (352). Affixal perceptual salience enables the recognition of an affix "as a unit of processing for lexical access" (Laudanna and Burani 1995: 360). A high perceptual salience of the affix favours morphological decomposition, which requires the use of information on the base and affix to access the word. The salience of the affix is dependent on factors such as orthography, which may lead to confusability[5] (Laudanna and Burani 1995; Laudanna et al. 1994), allomorphy, homonymy (Schreuder and Baayen 1995), frequency of the affix (Laudanna and Burani 1995; Plag and Baayen 2009), distribution of the affix (i.e. "co–occurrences with other relevant sub–lexical units" (Laudanna and Burani 1995: 352)), or affix productivity, which, according to Laudanna and Burani (1995), increment the bias towards morphological parsing. The factors cited in these studies do not constitute a complete list of all the conditions considered in the processing of complex words. They represent the awareness of science towards the complexity of word processing. Recent proposals based on experiments provide evidence of the inadequacy of sublexical and supralexical models, as well as dual–route models designed as a 'horse race', as demonstrated by Pollatsek et al. (2008).

Kuperman et al. (2008, 2009, 2010) proposed a multiple–route model as an alternative. According to these authors (2010: 94), the model considers the mental lexicon as a "long–term memory storage for lexical information". They note that: "The ease of access, and generally of lexical processing, depends in part on the amounts of information carried by words, which are defined by

5 According to Kuperman et al. (2010: 87), confusability refers to "[...] the ratio of word types in which the character string functions as a suffix and all word types ending in that character string [...]".

the accumulated knowledge of words and their paradigmatic and syntagmatic connectivity in the mental lexicon. The multiple–route model considers morphological structure as a conglomerate of sources of information, which contribute – to a different extent – to the recognition of polymorphemic words". The multiple–route model agrees with interactive dual–route models (Baayen and Schreuder 2000) that different sources of information may contribute to each other. As Kuperman et al. (2010: 94) present it, this model is a "multiple–route model of morphological processing, which considers morphemes, combinations of morphemes, and morphological paradigms and structurally complex words as sources of morphological information".

3.1 Combinations of affixes, morphological complexity of patterns, expectedness, semantic transparency and pattern character

There are several factors involved in determining how complex words are processed. In this section we devote our attention to some specific factors, such as the combination of affixes, morphological complexity of word–formation patterns, semantic transparency, expectedness, pattern character and the relations between them.

Models that refer to morphological decomposition as the process of accessing complex words imply that there is storage of morphemes (Taft 1981; Marslen–Wilson et al. 1994). Linguistic theories disagree as to whether morphemes are listemes, i.e., if they are stored as lexical units (Halle 1973, Lieber 1980 and Selkirk 1982), or if they merely actualize phonological rules (Beard 1995) or pertain to schemata (Booij 2010). Tomasello (2000: 238) emphasized that the acquisition of a particular language is founded on the mental storage of representations of concrete linguistic phenomena that suffer a gradual process of abstraction that culminates in the construction of schemata. Whether we take morphemes as listemes or as realizations of abstract representations, there must be information in memory that the stimulus may match with, so that the recognition of the word, or of the pattern of the word, takes place. The question remains whether storage prerequisites for access to a complex word are merely storage of those elements, or of the schemata or rules where those morphemes operate, or if more complex representations are required, i.e., the possible combinations between word formation rules or particular schemata, so that the stimulus may be recognized in accordance with data stored in the mental lexicon.

According to our study, it is not enough to have stored morphological constituents or schemata that are formally actualized by those constituents. Even if the representations of morphological constituents can obviate the access to the representation of the complex word, as demonstrated by Taft (2003); Taft and Kougious (2004); Libben (1998); and Schreuder and Baayen (1995, 1997), this is not enough to ensure access to the complex word. There must be stored information in the mind about the combinations between those morphemes or between the rules or schemata where they operate. This information is needed so that the affix combination as a whole may have a pattern character.

Caramazza et al. (1998), Schreuder and Baayen (1995), Frauenfelder and Schreuder (1991) and Ji et al. (2011), demonstrated that the processing of polymorphemic words requires not only decomposing them into their morphological constituents but also a more integrative operation. There must be an operation of integration of the constituents in the whole. Semantic transparency has a role in this operation. According to Libben et al. (2003: 51), semantic transparency "is crucial to the understanding of the manner in which multimorphemic words are represented in the mind". Studies such as Laudanna and Burani (1995), McQueen and Cutler (1998), Marslen–Wilson et al. (1994), Schreuder and Baayen (1995), Libben (1995; 1998; 2010), Libben et al. (2003), Bell and Schäfer (2013; 2016) and Frisson et al. (2008) debate the importance of semantic transparency in the way morphologically complex words are represented and accessed in the mental lexicon. This has been a matter of debate for compounds, specifically opaque compounds, because their meaning is non–compositional (Libben 1995). Nevertheless, semantic transparency is also important in affixed words (Taft 2004; Taft and Ardasinski 2006), as affixation may lead to non–regularities of meaning (e.g. Corbin 1987).

The fact that the semantics of a complex word may be non–compositional – that is, that it does not result from the semantic structures of its constituents – may be exemplified with Portuguese compounds. In Portuguese, *caravela–portuguesa* 'Portuguese man o' war' (*Physalia physalis*) designates a marine cnidarian, although its constituents are *caravela* 'caravel' and *Portuguesa* (feminine of 'Portuguese'). *Amor–perfeito* 'pansy' designates a flower, even if its constituents are *amor* 'love' and *perfeito* 'perfect'. The same occurs with affixed words. The word *Pré–Rafaelitas* 'Pre–Raphaelites' designates a group of artists, not located at a time before Raphael, and the verb *espalhar* 'to spread', constituted by the prefix *es–* and the nominal base *palha* 'straw', does not mean any event related to this noun. Words like these are characterized by a non–compositional meaning and are semantically opaque in the relation between their semantic structures and the semantic structures of their morphological constituents.

In contrast, words such as *espanholizar* 'to make something/someone Spanish' and *verbalização* 'verbalization' are compositional in meaning. The suffix *–iz–*, which permits the formation of denominal and deadjectival verbs (Pereira 2008; 2016), is added to the base *espanhol* 'Spanish'. The suffix –ção, which is a nominalizer (Rodrigues 2008; Rodrigues 2016), occurs with the verbal base *verbalizar*, in the theme form (*verbaliza–*)). Their meanings are semantically transparent. The question is whether every complex word with compositional meaning is easily interpreted by the speaker – or not. Words such as *comercializabilidade* are compositional in meaning, but they do not result in a clear interpretation for the speaker. Our study demonstrates that this is due to the non–expectedness of the combination of affixes involved in these kinds of words. To understand this, we need to reexamine the relationship between semantic transparency or compositionality and interpretability.

According to Plag (2003: 46), semantic transparency exists in words when "[...] their meaning is predictable on the basis of the word–formation rule

according to which they have been formed." This is a linguistic definition of semantic transparency. It holds true in the case of words such as *espanholizar* and *verbalização*.

Zwitserlood (1994: 344), analysing compounds, stated that "[t]he meaning of a fully transparent compound is synchronically related to the meaning of its composite words [...]." This is also the vision of Marslen–Wilson et al. (1994: 5), who indicate that "A morphologically complex word is semantically transparent if its meaning is synchronically compositional. Words like *happiness* and *unhappy* are semantically transparent because their meaning is directly derivable from the meaning of their stem {happy} together with their respective affixes {–ness} and {un–}. It is implausible that the lexical entries for words like this should not be related, in some way, to the lexical entry for the stem *happy*. This is the case of the compound *garça–branca* 'great egret (*egretta alba* or *ardea alba*)' and the suffixed noun *avaliação* 'evaluation'.

Libben (1998) proposes a representation of compounds at three levels: the stimulus, lexical and conceptual levels, and a double representation of semantic transparency at the conceptual level. One representation considers "the semantic relationship between the meaning of a morpheme within a compound and the independent meaning of that same morpheme" (Libben 1998: 37). The other representation deals with the transparency of the "compound as a whole" (Libben 1998: 37–38). Libben called this componentiality, which is lacking in *bahuvrihi* compounds, as pointed out by Libben (1998: 38). In these compounds the meaning is opaque, not because of the opacity of the constituents, but because the meaning of the compound is noncomponential. This is the case of *rabirruivo* 'redstart (*phoenicurus*)', from *rabo* 'tail' and *ruivo* 'red'. According to Libben (1998: 38–39), semantic transparency "is not monolithic but rather is composed of two broad types yielding eight possible transparency profiles [...]". This is in accordance with Schreuder and Baayen (1995: 140) who stipulate that "[...] a semantically transparent relation between a complex word and its constituents can be modelled as a substantial overlap between the set of (semantic) representations of the complex word and the sets of representations of its constituents".

This overlap is also a property of semantic transparency of affixed words. Nevertheless, semantic transparency encounters different problems in compounds and affixed words. As pointed out by Libben (1994), this is because the analysability of compounds contrasts with the open list of roots, whilst the affixed words must be matched with constituents that belong to a closed list. Further, affixes do not have referential semantics, which may make their identification as units difficult.

Libben et al.'s (2003: 62) definition of semantic transparency of compounds, as "the manner in which a morpheme's semantic characteristics in a multimorphemic word correspond to its semantic characteristics as a free–standing lexical item" does not apply to affixed words, as morphemes do not stand as free lexical units. Even in psycholinguistic studies, the definition of semantic transparency is grounded in linguistic structures. Semantic opaque-

ness corresponds to non–compositionality and semantic transparency with compositionality.

However, in attempting to define semantic transparency based on the analysis of the word constituents, a problem arises with some words. Although their meaning results from the combination of the meanings of constituents, their semantics is not easily interpretable to the speaker. This is the case of potential constructs such as *anti–anti–anti–anti–nacionalismo* 'anti–anti–anti–anti–nationalism'. The meaning of the construct is compositional, but the speaker has to make an explicit effort to count the number of *anti–* and apply it to the meaning of each one of the constructs (*anti–anti–nacionalismo* is the 'attitude of being against the anti–nacionalismo'; *anti–anti–anti–nacionalismo* is the 'attitude of being against the attitude of being against the anti–nacionalismo'; *anti–anti–anti–anti–nacionalismo* is the 'attitude of being against the attitude of being against the attitude of being against the anti–nacionalismo'). The semantics of *anti–anti–nacionalismo* seems to be achieved effortlessly. However, the subsequent constructions require the speaker to think explicitly about the compositionality involved in the constructs.

This means that predictability, anchored in each of the word formation rules or schemata involved in each word, although supporting compositionality of the meaning, may not lead to interpretability from the speaker's point of view. One of the reasons is that there may not be a mental pattern underlying the concrete word in the speaker's mind, if we think of approaches such as Tomasello (2000: 238), which highlights the construction of abstractions from concrete linguistic phenomena represented in the mind. This is the case for compounds, as pointed out by Bell and Schäfer (2013: 1). It is also true for affixed words, whose combinations may not result in a pattern recognizable by the speaker. Thus, we need to bear in mind *interpretability*, which depends on the conjunction between compositionality and the experience of the speaker. Compositionality is defined in linguistic terms. It indicates that the meaning of the complex word results from the meanings of the morphemes that constitute it. Where there is a lack of compositionality, meaning is opaque. Interpretability depends on data stored in the speaker's mind. Because it is a psycholinguistic factor rather than a structural one, interpretability may vary from speaker to speaker, depending on variables that come from the experience of the individual. This experience is dependent on the individual being exposed to stimuli and varies according to diachronic, social, dialectal and idiosyncratic factors that determine language usage.

Compositionality of meaning is in accordance with predictability of the word formation rules or schemata (linguistic level). Interpretability reflects the ease of meaning processing without the need to explicitly think of its constituency, based on the experience of the speaker (psycholinguistic level). What becomes salient in our study is that, for the processing of a suffixed word, it is not enough that a word consists of a compositional meaning. The interpretability of the word is dependent on the experience that the speaker has towards not only the affixes constituting the word, but also the combination of the involved suffixes. This is so that the word will be expected by him/her.

This is in line with Bell and Schäfer (2016: 163) who, in advancing their evidence from earlier work (2013), stipulate that "the semantic transparency of a compound is a function not only of the transparency of its constituents but also of the semantic relation between them". The authors "[...] test the hypothesis that perceived transparency is correlated both with the expectedness of the constituents themselves and with the expectedness of this relation". (163) Semantic transparency is viewed as a scalar notion that combines predictability and meaning relatedness (Bell and Schäfer 2016).

In the case of affixed words, the experiments with *–bil/–al–izabilidade* words show that the expectedness of the combination between affixes is crucial for the semantic interpretation of the word. Although *–bil/–al–izabilidade* words are constituted by components that adhere to structural constraints, and have compositional meanings, the combination of the affixes present is not expected. That combination is not frequent, and thus, the semantics of these words become opaque in the speaker's mind. As Baayen (2007: 84) notes, "it must be advantageous for the brain to keep track of detailed combinatorial probabilities". The affix combination in *–bil/–al–izabilidade* words is not a probable one, as results of this study show.

In their analysis of compound words, Bell and Schäfer (2016: 157) state that there is a relation between frequency and expectedness, productivity and expectedness, and expectedness and semantic transparency. They emphasize, "By expectedness, we essentially mean familiarity and frequency of use. It therefore seems plausible that greater expectedness will lead to greater ease of processing, as greater familiarity with an item or relational structure will mean that the language user has had more practice with it" (195). They add (in their footnote 10) that "There could be also a positive feedback effect, whereby users are more likely to choose to use items they themselves would find easier to process, so that ease of processing conversely leads to greater frequency of use". Our study is in line with their proposal, as *–bil/–al–izabilidade* words manifest a low type frequency, which points to a non–productive affixal chain. Although the meaning of these words is compositional, the affix combination is not expected, which, besides the extension of the affixal chain, hinders the processing of the words containing it. The fact that the affixal combination is not expected implies that it does not correspond to a word–formation pattern of contemporary Portuguese, as Experiment 2 suggests.

The importance of the particularities of each language system for word processing are underlined by Laudanna and Burani (1995: 347): "If the distributional and linguistic parameters of derivational morphology are relevant in shaping the organization of the mental lexicon and, in particular, the mechanisms which allow access to lexical representations, then it follows that as the range of values of those parameters vary from one language to another, the more they may lead to varying organizations of the lexical access systems in the different languages (or languages types)". This converges with Sims and Parker (2015), who specifically argue on the importance of language–specific properties to the affix combination.

Apart from differences between languages, we also need to bear in mind the particularities of the language usage of each speaker that gives rise to his/her mental lexicon.

The combination of affixes that are constituents of complex words is a crucial factor in the processing of the word. If a specific affix combination is frequent, and thus, expected by the speaker, it facilitates interpretability. The combination of affixes must be ruled by structural constraints (section 2), but the processing of the words that contain it is dependent on variables. These variables are frequency of combination, expectedness and semantic transparency plus interpretability. Joined together, they act upon the pattern character of the affixal combination, which we have defined on section 1.

4. Structural constraints, affixal combinations and patterns

Word–formation is characterized by regularities and semi–regularities in semantic, phonological, syntactic and morphological levels that have been captured by linguists through different models (e.g. word formation rules (Halle 1973; Aronoff 1976; Corbin 1987) and schemata (Booij 2010)). It is not our aim to develop those models here; cf. Baeskow (2015) for that purpose. Rules or schemata formulated by linguists state explicit mental patterns that enable the speaker to form and analyse words that are semantically and formally predictable or half–predictable (Plag 1999; Bauer 2001; 2005; Gaeta and Ricca 2015).

Throughout the history of a specific language, the productivity of a pattern, that is, the degree to which a pattern may be used to the construction of new lexemes, may change (cf. Scherer 2015 and Rainer 2015). In a particular state of that language, productive patterns allow for the construction of words. In a different state of the same language, those patterns may be no longer productive, but they are still patterns, if they allow the speaker to interpret and use the word, by means of correlations manifested with other words (see words such as *depth* illustrated below).

The concept of productivity is a complex one, as demonstrated by the many different theories pertaining to it, which are in turn linked to the different features of word–formation associated with the concept (cf. Rayner 1987; Bauer 2001; 2005; Gaeta and Ricca 2015 on the several theories of productivity and its complexity). Productivity, as Corbin (1987) stipulated it, involves two different aspects: profitability (the number of words producible by means of that pattern) and availability (the possibility of the speaker to use the pattern to form new words), *rentabilité* and *disponibilité*, as coined by Corbin (1987) (cf. Carstairs–McCarthy 1992).

The distinction between pattern and productivity of the pattern is important due to diachronic changes and processing consequences. This is why we deal in this work with the concept of pattern character, which should be kept apart from the concept of productivity. To illustrate changes in productivity, Scherer (2015) points out the English suffix *–th*, which was productive in Old and Middle English to form deadjectival quality nouns and is today unproduc-

tive. This means that in contemporary English it is not possible to form quality nouns using the suffix *–th*. However, quality nouns such as *breadth*, *length*, *depth*, *truth*, *strength*, *warmth* and *width* and the adjectives *broad*, *long*, *deep*, *true*, *string*, *warm* and *wide* constitute a pattern, organised by means of the suffix *–th*, the semantics of the words and their lexical category. These kinds of pattern have been described in literature through redundancy rules or word–structure rules (Plag 2003: 36–37). As the suffix *–th*, the affixal chain under analysis in this study is not productive. Nevertheless, the suffix *–th* possesses a role on the aggregation of a pattern. We are trying to verify here if the affixal chain *–bil/–al–iza–bil–idade* may function as a mental pattern, having enough strength to enable the speaker/listener to interpret/use the words containing it.

Apart from processing conditions, productivity is ruled by structural constraints that restrict the combinations between affixes and bases in each one of the patterns, and the bases and phenomena that operate in patterns lacking affixation (e.g. conversion) or working with other kinds of morphological derivational devices (backformation; compounding) (cf. Rainer 2005; Rodrigues 2009; 2014; 2015). Structural constraints are described as (in)compatibilities working on the phonological, semantic and syntactic structures, in different tiers of each of the structures. Structural constraints predict that words such as *generalizabilidade* 'generalizability', *legalizabilidade* 'legalizability', *localizabilidade* 'localizability', and *materializabilidade* 'materializability' are possible forms. In fact, they are the result of multiple affixation that manifests a series of derivational rules. This multiplicity is illustrated in Table 1.

Pattern/ Combination 1	Pattern/ Combination 2	Pattern/ Combination 3	??Pattern??/ Combination 4
matéria$_{N}$> *materi–*$_{N}$*–al*$_{ADJ}$	*material*$_{ADJ}$> *material*$_{ADJ}$*–iz–*$_{V}$	*materializ–*$_{V}$> *materializ*(*a*)$_{V}$*–vel*$_{ADJ}$	*materializ*(*a*)$_{V}$*–vel*$_{ADJ}$> *materializabil–*$_{ADJ}$*–idade*$_{N}$
comércio$_{N}$> *comerci–*$_{N}$*–al*$_{ADJ}$	*comercial*$_{ADJ}$> *comercial*$_{ADJ}$*–iz–*$_{V}$	*comercializ–*$_{V}$> *comercializ*(*a*)$_{V}$*–vel*$_{ADJ}$	*comercializ*(*a*)$_{V}$*–vel*$_{ADJ}$> *comercializabil*$_{ADJ}$*–idade*$_{N}$
conta–$_{V}$> *conta–*$_{V}$*–vel*$_{ADJ}$	*contável*$_{ADJ}$ > *contabil–*$_{ADJ}$*– iz–*$_{V}$	*contabiliz–*$_{V}$> *contabiliz*(*a*)*–*$_{V}$*–vel*$_{ADJ}$	*contabiliz*(*a*)*vel*$_{ADJ}$> *contabilizabil–*$_{ADJ}$*–idade*$_{N}$

Table 1. Patterns and Combinations between affixal derivational operations

In Table 1, there are different operations of word–formation that operate by affixation. Four remarks about this must be made before we proceed with the description of data.

First remark: In the first three columns, we placed the words *Pattern* and *Combination*. In the last, we used the words *Combination* and *Pattern* surrounded by question marks. By *Combination* we mean that the complex

word manifests a combination of morphological constituents. This does not mean that we follow an Item and Arrangement model (e.g. Bloomfield 1933). It implies that the word that results from that formation shows a certain combination of morphological constituents. By *Pattern* we mean that that specific mechanism reflects data in the speaker's mind, which enables him/her to productively form and/or recognize words that obey the operations involved in that formation. Combinations and patterns are not obligatory requisites to word–formation. Conversion, for instance, is a word–formation procedure that lacks combination, because it does not operate with the addition of morphemes to a base. Nevertheless, it is a productive pattern as long as it functions as a productive way to form words and to enable their recognition (cf. Bauer and Valera (2005) Valera (2015); Rodrigues (2001; 2009; 2013)). This is the case of conversion of noun to verb in English, such as *hammer*N/*to hammer*V. In contrast, some creative ways to form words and compounds do not follow a pattern that exists in the speaker's mind, although they may use the combination of affixes and/or of bases. This is the case of compounds such as Portuguese *cabra–cega* 'blind man's buff', which results from the combination of *cabra* 'goat' and *cega* 'feminine of blind', and of affixed words such as *–bil/al–izabilidade* words, as the experiments in this work demonstrate. As one of the anonymous reviewers pointed out, the essential difference between words such as *cabra–cega* and *–bil/al–izabilidade* words consists in the fact that *cabra–cega* has an idiosyncratic meaning that does not actualise a specific morphological pattern, contrarily to *–bil/al–izabilidade* words. We think that, although the meaning of *–bil/al–izabilidade* words depends on structures that inform several morphological patterns that are involved in the formation of *–bil/al–izabilidade* words, the whole, which is formed by the set of patterns, may not constitute a pattern on its own. Bearing in mind that one of the aims of this study is to determine whether *–bil/–al–izabilidade* words constitute a derivational pattern in contemporary Portuguese, we leave the column with the question marks adjoining the word *Pattern*.

Second remark: The formal representation in Table 1 does not imply that a concatenative account of word formation, grounded on an item and arrangement model (cf. Bloomfield 1933), is the theoretical approach adopted by us. This merely represents the position of suffixes and the patterns they work in. Rather, we follow a mental, paradigmatic approach, assuming that in the mental lexicon there is no need for the real existence of a specific word so that a lexeme may be constructed upon it as a base, as long as patterns exist in the mind (cf. Booij 2010). This is not the same as assuming that patterns in the mind dispense with fully–spelled out words which instantiate that pattern, which is in accordance with approaches such as Burzio's (1991 [1995]; 2002) and Blevins (2016), brought to our attention by one of the anonymous reviewers.

Third remark: When a hyphen is provided, it indicates that the form of the lexeme or affix is not able to occur as a word in syntax, that is, as a surface form. For instance, the root *materi–* does not occur as a word in syntax, whilst *matéria* does.

Fourth remark: Suffix *–vel* is characterized by allomorphy: when appearing at the end of the word, its form is *–vel*; in the middle of the word, it is *–bil–*. Diachronic reasons underlie this difference, but synchronic explanations may be stated: i) position of the suffix (*–vel* if it is the last morpheme of the word (*permeável* 'permeable'), excluding inflection; *–bil–* if it is not the last derivational morpheme of the word (*permeabilidade* 'permeability')); and ii) activity or inactivity of the suffix to the word formation pattern (*–vel* if it is active, that is, if the affix plays a role in the formation of the word; *–bil–* if it is inactive, that is, if the affix does not work on the morphological pattern that constructs the word). According to the second explanation, the affix *–vel* is active, i.e., it works on the formation of the adjective *permeável*, from the verb *permear* 'to permeate'; but it does not work on the formation of the noun *permeabilidade*, although the affix is contained in it. The formation of *permeabilidade* occurs by means of the affix *–idade* and not of the affix *–vel*/*–bil–*, which is not responsible for the construction of the quality noun. Because this allomorphy also appears in basic words, the first explanation is more accurate than the second, unless we consider that *–vel*/*–bil–* is always interpretable with the semantics of 'that is prone to something', even when we are dealing with words such as *sensível* 'sensitive' and *legível* 'legible', which are not constructed in Portuguese. This sheds light on the hypothesis that a pattern may be supported by words that are not constructed in a certain language, but that have a complex morphological structure, in the sense that at least one morpheme can be analysed as such. This morpheme may be active in a word and inactive in another one, but the preservation of semantic correspondence strengthens its pattern character.

The affix combinations shown in Table 1 that correspond to productive patterns of contemporary Portuguese (combinations 1–3) are represented in Tables 2a, 2b, 3 and 4. In Table 5 there is a combination that permits the construction of quality nouns in *–idade* from adjectives ending in *–vel*. This does not represent the totality of combinations posted in combination 4 of Table 1, because an aim of this study is to establish if this last combination corresponds to a pattern.

base	suffix	Derivative	Semantics of the derivative
N (root)	*–al*	ADJ	'that has a relation with Nb'
comércio 'commerce' (*comerci–*) *matéria* 'matter' (*materi–*) *fisco* 'revenue' (*fisc–*)		*comercial* 'commercial' *material* 'material' *fiscal* 'fiscal'	

Table 2a. Combination of morphemes that correspond to Pattern of construction of relational denominal adjectives ending in *–al*

base	suffix	Derivative	Semantics of the derivative
V (past participle theme)	*–vel*	ADJ	'that is prone to be Vpast participle'
beber 'to drink' (*bebi–*) *arar* 'to plough' (*ara–*) *partir* 'to break' (*parti–*)		*bebível* 'drinkable' *arável* 'ploughable' *partível* 'breakable'	

Table 2b. Combination of morphemes that correspond to Pattern of construction of deverbal adjectives in *–vel*

The combination shown in Table 2a corresponds to a productive pattern in contemporary Portuguese. It forms adjectives meaning 'that has a relation with N' from nouns (*e.g. carne* 'flesh', *ambiente* 'environment', *semana* 'week', *floresta* 'forest', *Provença* 'Provence', *dente* 'tooth' and *organização* 'organization' are the respective deriving words of *carnal* 'carnal', *ambiental* 'environmental', *semanal* 'weekly', *florestal* 'of forest', *provençal* 'from Provence', *dental* 'dental' and *organizacional* 'organizational'). The pattern shown in Table 2b is repeated in Table 4 and shown in that context.

base	suffix	Derivative	Semantics of the derivative
Adj (root)	*–iz–*	V	'to turn something into ADJb'
comercial (*comercial*) *material* (*material*) *fiscal* (*fiscal*)		*comercializ*(*ar*) *materializ*(*ar*) *fiscaliz*(*ar*)	

Table 3. Combination of morphemes that correspond to Pattern of the construction of verbs in *–iz–* from adjectives ending in *–al*

The combination of derivational morphemes in Table 3 constitutes a productive pattern in contemporary Portuguese. It is inscribed in a major pattern according to which a simple or complex noun or adjective constitute the bases of verbs. Simple bases, not constituted by more than one morpheme, may be illustrated by *jardim* 'garden' and *suave* 'smooth, soft', which are the respective bases of the verbs *jardinizar* 'to garden' and *suavizar* 'to soften'. Suffix *–iz–* may attach to complex bases with different suffixes, as illustrated by words such as *infantilizar* 'to infantilize', and *americanizar* 'to Americanize' from the denominal adjectives *infantil* 'childish' (*infant–il*), *americano* (*Améric–ano*) 'American' (cf. Pereira 2016 for an exhaustive description of the pattern). It is important to note that the *–ar* ending of the verb is constituted by the theme vowel (*–a–* in 1st conjugation verbs) and the infinitive morpheme (*–r*). The impersonal infinitive form corresponds to the citation form in Portuguese tradition and has no derivational role in the verb construction (Rodrigues 2016a).

base	suffix	Derivative	Semantics of the derivative
V (present theme)		ADJ	
comercializar 'to commercialize' (*comercializa–*) *materializar* 'to materialize' (*materializa–*) *fiscalizar* 'to supervise' (*fiscaliza–*)	*–vel*	*comercializável* 'marketable' *materializável* 'that is prone to be materialized' *fiscalizável* 'that is prone to be supervised'	'that is prone to be Vpast participle'

Table 4. Combination of morphemes that correspond to Pattern of construction of adjectives in *–vel* from verbs in *–iz–*

Table 4 shows a combination that permits Portuguese speakers to construct and recognize *–vel* ending adjectives from verbs in *–iz–*. It is a productive pattern in contemporary Portuguese. This pattern is inscribed in a more general pattern, according to which suffix *–vel* may attach to verbs constituted by several morphological structures (simple verbs, such as *beber* 'to drink' and *palpar* 'to touch, to palpate', which are the bases of *bebível* 'drinkable', *palpável* 'palpable', shown in Table 3b); complex verbs such as *enternecer* 'to move someone', *danificar* 'to damage', *esventrar* 'to disembowel', which are, respectively, the bases of *enternecível* 'that is prone to be moved', *danificável* 'that is prone to be damaged' and *esventrável* 'that is prone to be disemboweled', among many others) (Rodrigues 2016b).

base	suffix	Derivative	Semantics of the derivative
ADJ (root)		N	
razoável 'reasonable' (*razoabil–*) *rendível* 'rentable'(*rendibil–*) *permeável* 'permeable' (*permeabil–*)	*–idade*	*razoabilidade* 'reasonability' *rendibilidade* 'rentability' *permeabilidade* 'permeability'	'quality of ADJb'

Table 5. Pattern of construction of nouns ending in *–idade* from adjectives ending in *–vel*

Table 5 shows the pattern by which quality nouns with the suffix *–idade* have bases of adjectives ending with the suffix *–vel*. This pattern inscribes itself in a major pattern of the formation of that kind of nouns because it is possible to construct *–idade* nouns from adjectives with different morphological structures. Simple bases such as *ameno* 'pleasant' and *atroz* 'atrocious' are the

deriving words of *amenidade* 'pleasantness' and *atrocidade* 'atrocity'. Complex words with suffixes other than *–vel* may also function as deriving words of *–idade* nouns. This is the case of *espiritualidade* 'spirituality', from the adjective *espiritual* 'spiritual', formed from *espírito* 'spirit' by means of the suffix *–al*, *oleosidade* 'quality of being oily', from the adjective *oleoso* 'oily', formed from óleo 'oil' with the suffix *–oso*, and *cristalinidade* 'crystallinity', from the adjective *cristalino* 'crystalline' formed from the noun *cristal* 'crystal' with the suffix *–ino*. Rio–Torto (2016) presents a systematic description of this pattern.

4.1 Correspondence of affixal combinations with patterns

Tables 2–4 purposefully contain the words observed in Combinations 1–3 of Table 1. In contrast, Table 5 contains words other than the ones inscribed in the fourth column of Table 1. This is because the formation of words by means of the particular affixes operating in patterns 3–5 (*materi–al–iz*(á)*–vel*, *comerci–al–iz*(á)*–vel*) is productive in contemporary Portuguese. However, Table 5 shows us that what is productive in terms of formation of *–idade* nouns is its formation from adjectives in general, and not specifically the formation of nouns from adjectives that manifest the structure that we want to focus on: base*–bil*/*–al–iz–bil–idade*.

Because this specific combination has low frequency, we need to verify if it corresponds to a word–formation pattern or not by means of Experiments 1 and 2. Correspondence to a word–formation pattern may be judged by the presence of the following conditions:

a) Structural conditions:
   i) Correspondence of morphological devices to word–formation devices of Portuguese;
b) Processing conditions:
   i) Interpretation, by native speakers, of any word formed with that combination of affixes, as long as they know the semantics of the root and as long as they have experience of the combination of affixes involved, so that the combination is expected by the speaker, as established in section 3.1;
   ii) Ability of the speaker to use that device to form words.

For instance, if a native speaker hears/reads the word *geringoncista* (coined by Rui Pereira, politician, http://www.publico.pt/politica/noticia/eu–geringoncista–1724234, accessed on the 25th of February, 2016) for the first time, he/she may not specifically know the particular conjuncture that had aroused its coinage. He/she will need to read the text to understand it. Nonetheless, as the base *geringonça* 'contraption' and the suffix *–ista* are recognized in Portuguese, and as this formation actualizes an expected (condition b.i) derivational device (condition a.i), the native speaker is able to interpret *geringoncista* (condition b.i) as 'someone with a strong connection with *geringonça*', 'someone that makes or supports *geringonças*', independently of what this base refers to. (In this particular case, *geringonça* informally applies to the agreement made by left–wing parties and the centre–left wing party after the elections for

the 13th Portuguese legislature (October, 2015).) The coinage of *geringoncista* is enabled by condition b.ii).

Combinatoriality of affixes must be checked at two levels:

i) One–to–one combination: word–formation working in these patterns is possible, as we may have adjectives in *–al* from nouns (*global*), verbs in *–iz–* from adjectives in *–al* (*comercializar*), adjectives in *–vel* from verbs in *–iz–* (*espanholizável*), and nouns in *–idade* from adjectives in *–vel* (*dilatabilidade*).

ii) Multiple combination of all patterns: until Pattern 3, the combination is possible and frequent (*glob–al–iz(á)–vel*). Combination 4 must be checked according to the conditions we stated for the existence of a word–formation pattern.

The one–to–one combination indicates that, theoretically, there could be nouns in *–idade* whose bases contain the sequence *–bil/–al–iz(a)–bil–*, as there are no structural constraints on the combination between these particular affixes (cf. Bauer, Lieber and Plag 2013: 494), which are defined according to phonological, morphological, syntactic and semantic properties of the affixes and bases. However, we must question how far syntagmatic combinations of suffixes maintain semanticality, i.e. semantic well–formedness of the constructed word, (Pustejovsky 1995) and grammaticality. The one–to–one combination may function as a pattern, being expected by the speaker's mind, but the presence in the same word of all the combinations may not be a pattern, if it is not expected, or, differently stated, if the set of the combinations as a whole does not correspond to stored information in the mind.

Despite the good one–to–one combination of the suffixes *–al/–iz–*, *–iz–/–vel* and *–bil–/–idade* and the structural well–formedness of the derivatives that result from those patterns, *–bil/–al–izabilidade* words have a low frequency in Portuguese. This means that the combination *per se*, and not the specific word containing it, is not expected by the speaker, which possibly shows the low probability of its constitution as a word–formation pattern. The possible hypothesis that this occurs because there are more than three suffixes must be treated with caution. The same low frequency affects words such as *generalizabilidade* 'generalizability', *legalizabilidade* 'quality of being legalizable', *localizabilidade* 'localizability', where the segment *–al* does not correspond to a suffix in Portuguese or where Pattern 1 (N> *–al* ADJ; V> *–vel* ADJ) does not occur. In *generalizabilidade*, *legalizabilidade,* the segment *–al* may function inside Pattern 1, as *legal* 'legal' and *geral* 'general' are adjectives where *–al* is recognizable, not as a derivational morpheme, but through semantics and syntax. This occurs despite the lack of a noun deriving them or the allomorphy occurring between *general–* (allomorph of the adjective *geral* 'general' not occurring in Portuguese in the form of the adjective) and *geral* 'general' (occurring in Portuguese as a form of the adjective) and *legal* 'legal' and *lei* 'law'. In the case of *localizabilidade*, *local* 'local' is a noun and not an adjective. In contrast, as will be observed in our study, words bearing multiple pleonastic suffixation function differently when compared with these. Nevertheless, despite the morphological differences between *generalizabilidade*/*legalizabilidade*

and, say, *comercializabilidade*, where the segment *–al* corresponds to an active morpheme, the low frequency of both kinds of words indicates the low type frequency of the affix combination, as a whole.

5. Experiments

5.1 Methodology

Because our main objective is to understand if *–bil/–al–izabilidade* affixal chain constitutes a word–formation pattern of contemporary Portuguese or if the few *–bil/–al–izabilidade* words are not instances of a word–formation pattern, we evaluated their processing by native speakers using an acceptability judgment task and a recall task. In the two experiments, the processing of those words was compared with the processing of evaluative words with pleonastic suffixation. The acceptability judgment task is not enough to give us the status of the word–formation device as a pattern. A production task, such as a recall task, is needed (cf. Kuperman et al. (2010) for other limitations of an acceptability judgment task.)

5.1.1 Participants

Group A (participating in the acceptability judgment task) comprised 22 (11 men and 11 women) native speakers with normal to corrected–normal vision and hearing. Participants were undergraduate students of Instituto Politécnico de Bragança (Portugal) without knowledge of linguistics, who participated in this study voluntarily. A different group of 22 students (11 men and 11 women), Group B, participated in the recall task. This group met the same criteria as Group A but comprised different individuals. In both groups the age range was 18 to 22.

5.1.2 Stimuli

The same target words (Appendix A) were used in both tasks. In the recall task, the target words were inserted in sentences (Appendix B). Target words consisted of 76 words belonging to the following categories: 27 neutral words (these words are frequent words of Portuguese, basic and with no relation to the morphological chains under focus); 12 pseudo–words (created *ad hoc* for this investigation); 13 frequent evaluative suffixed words (evaluative nouns and adjectives with pleonastic suffixation that are frequent in Portuguese, according to corpora. Because in Portuguese multiple affixation in evaluative formation is not frequent, frequent words with evaluative affixation do not contain multiple affixation); 13 non–frequent or created evaluative nouns/adjectives with multiple suffixation; 10 *–bil/–al–izabilidade* ending words (As corpora indicate, *–bil/–al–izabilidade* words have a low frequency, we added other words with the same affixal combination to the words we found on Google. It was important to evaluate not whether a specific word exists, but if the pattern exists). Apart from functioning as a contrast to complex words, pseudo–words and neutral words were also intended to behave like fillers, to

avoid participants constructing their answers according to predictions based on regularities found among the stimuli.

All the target words were provided in their citational form (If the word was a verb, it was given in the impersonal infinitive form; if the word was a noun, in the singular form; if an adjective, in the singular masculine).

Frequencies of the target words are provided in Appendix A. A caveat must be taken in what regards frequency values. Frequency values correspond to the values displayed in Corpus de Referência do Português Contemporâneo (CRPC) (http://alfclul.clul.ul.pt/CQPweb/crpcfg16/), which contains 311.04 million words from written (309.8 million words) and spoken texts (1.6 million words). Only the written texts are available to be searched online (cf. http://alfclul.clul.ul.pt/CQPweb/doc/CRPC_Description_2.pdf). Hence, slang words such as *porra* (lit. 'cum') 'shit', very frequent in oral productions, emerge in formal written texts with low frequency, which may conceal data for readers who are not native Portuguese speakers. The same occurs with evaluative nouns such as *cafezinho*, *maçãzinha*, and all of those that we classified as frequent evaluative nouns. Evaluative nouns have a high usage in oral, informal productions in Portuguese. With the aim of highlighting that frequencies of CRPC do not reflect the oral and informal usage of Portuguese language, we decided to provide for each word the number of results available through Google. Even if Google displays only written texts, many texts have an informal and slang style, which may bias towards the oral usage of the language. A search in Google provides a considerable rate of occurrences (Appendix A) of the forms that are here classified as frequent. The discrepancy of frequencies indicated by CRPC and the number of results given by Google is evident, since CRPC are mainly built from formal texts. This discrepancy is well seen in the comparison between the frequency and Google results of the noun *camisa* 'shirt' (6.77/million *vs.* 175 000 000 results in Google) and the ones of *construção* 'construction' (231/million *vs.* 181 000 000 results in Google). Anyone can imagine that the first one has a much higher frequency usage than the last one in oral contexts. However, Google does not provide frequencies for Portuguese language and even the results obtained do not reflect the oral usage of language, as can be deduced from the results displayed for the word *cafezinho* 'diminutive of coffee' (only 8 630 000 results in Google), which is largely produced by Portuguese and Brazilian people. Regarding forms with the affix chain *–al–/–bil–iza–bil–idade*, they are not frequent at all. Google searches give back words containing the chain with a low rate.

The neutral words were constituted by 7 adjectives, 3 adverbs, 14 nouns, 2 verbs and 1 pronoun. There were 3 slang words. The minimum length of the words in letters was 3 and the maximum length of the words in letters was 13.

Pseudo–words presented syllable structures compatible with the ones of Portuguese. They did not contain phonological segments similar to those corresponding to derivational morphemes of Portuguese. The minimum length of pseudo–words in letters was 5 and the maximum length of the pseudo–words in letters was 10.

Frequent evaluative suffixed words were constituted by 10 nouns and 3 adjectives. Each of the words had one of the following evaluative suffixes: *–inho/a*, *–ito*, *–ino–*, and the sequence *–z–inho*. The minimum length of these words in letters was 7 and the maximum 10.

Non–frequent/created evaluative words were constituted by 11 nouns and 2 adjectives. Each one of the words contained the minimum of 2 evaluative diminutive suffixes (*–inha–zinha*; *–inhozinho*; *–inhozito*; *–acho–zinho*; *–iço–zinho*; *–ino–zinho*) and the maximum of 3 diminutive suffixes (*–inh–oto–zinho*; *–inh–ita–zinha*; *–in–inho–zinho*). The minimum length of words was 11 letters and the maximum length of the words in letters was 17.

Words containing the affixal chain *–bil/–al–izabilidade* were nouns. The minimum length of the words in letters was 15 and the maximum length of the words in letters was 22. Type frequency of the affixal chain *–bil/–al–izabilidade* is 0/million, according to Corpus de Referência do Português Contemporâneo (CRPC).

Sentences presented at the recall task contained the target words described above. Sentences had at the most two clauses, subject to the condition that one of them was a relative clause. Otherwise, sentences only contained one clause. Sentences might not contain more than one adjunct (adverbial phrase or prepositional phrase) of the predicator. Lexemes that do not constitute the target in the sentences were frequent or belong to a frequent morphological type. The target word did not appear as the last word of the sentence. The minimal length of sentences containing neutral words, in seconds, was 2.00 s; the maximum 4.82 s. The minimal length of sentences containing pseudo–words was 2.00 s; the maximum was 3.12 s. In sentences containing non–frequent evaluative words, the minimum length was 1.99 s; the maximum was 3.03 s. In sentences with frequent evaluative words, the minimum length was 1.82 s and the maximum was 3.38 s. In sentences with *–bil/–al–izabilidade* words, the minimum length was 2.76 s. and the maximum 4.17 s.

5.1.3 Procedure

5.1.3.1 Acceptability judgment task

Group A was given 76 words randomly listed on paper. The participants had to classify each of the 76 words as 'aceitável em português' ('acceptable in Portuguese') or as 'não aceitável em português' ('non–acceptable in Portuguese').

5.1.3.2 Recall task

Group B was presented with 76 verbal utterances, each containing one of the 76 words described as stimuli. The stimuli were presented in pre–recorded form spoken by a male native speaker of European Portuguese. The sentences were randomly listed. Individuals were asked to orally reproduce the sentences one–by–one immediately after hearing each one of the sentences. The recalled sentences were recorded. The participants' productions were classified into

seven types. Because our interest focused on the specific words and not on the entire utterance, the classification types refer only to the target words.

The seven types of productions obtained in the recall task included the following:

1. Total lexical match (TLM): Lexeme was exactly repeated.
2. Partial phonological mutation (PPM): Phonological structure was partially changed so that the produced form does not coincide with a Portuguese word.
3. Total semantic match (TSM): No repetition of the lexeme, but production of a lexeme semantically equivalent to the heard one. This may happen through the substitution of the lexical morpheme of the word by a synonymous one (e.g. *também* 'also' instead of *igualmente* 'also'), or through the preservation of the lexical morpheme and substitution of derivational morphemes by synonymous ones (e.g. *cadelinha* instead of *cadelinhazinha*; *lapinhozinho* instead of *lapinhozito*). In this last case, the meaning of the word is not changed, as *–ito* and *–inho* share the same meaning; both affixes operate in the construction of evaluative–diminutive words and are semantically equivalent.
4. Partial semantic match (PSM): No repetition of the same lexeme, but production of a lexeme only partially equivalent to the heard one. Partial equivalence is in this case only semantic or semantic and formal. This may happen through the substitution of the lexical morpheme of the word (e.g. *espacinho* 'small space' instead of *casinhotozinho* 'very small house'), or through the preservation of the lexical morpheme and substitution or addition of derivational morphemes by others (e.g. *aranh**ão**zinho* 'small spider' instead of *aranh**iço**zinho* 'very small spider'; *sociabiliza**ção*** 'event of turning someone/becoming sociable' instead of *sociabiliza**bilidade*** 'quality of being able to socialize').
5. Lexical category mismatch (LCM): Phonologically the lexeme was correctly recalled, and there is a phonological match between the stimulus and the produced word. However, it was inserted in a syntagmatic context proper of a lexical category that is different from the one of the stimulus (e.g. **Era riachozinho quando as pessoas se banhavam no verão*. '*It was small rivulet, when people bathed in the summer.' instead of *Era um riachozinho onde as pessoas se banhavam no verão*. 'It was a small rivulet where people took bath in the summer'). In this situation, *riachozinho* 'rivulet' was recalled as an adverb, instead of as a noun, although in Portuguese the word does not occur as adverb.
6. Other real word: An existing lexeme was produced, but bearing no relation to the semantic or phonological structures of the heard one (e.g. *Terra* 'Earth' instead of *céu* 'sky').
7. No answer (NA).

5.2 Predictions

Bearing in mind the correlations between frequency, semantic transparency, affixal salience, combination between affixes, morphological complexity of word–formation patterns, expectedness and pattern character, constructed upon the theoretical background discussed in sections 3.1 and 4 and that semantic transparency may be judged by total semantic match, partial semantic match, or total lexical match (opposed to partial phonological mutation), we make the following predictions:

1. Bearing in mind the correlation between frequency and transparency demonstrated by Bell and Schäfer (2016), a complex word, specifically in the case under study, a complex word with multiple derivational affixation, will be perceived as less transparent the lower the type frequency of the affixal chain. The lower the type frequency of the affixal chain is, the less expected it is. This means that the correlation between transparency and expectedness is not only observable in specific words, that is, the tokens, but it extends to the combinations of affixes as types (Hay and Baayen 2002). This implies that:
2. A complex word will be perceived as more transparent the more expected the combination between semantic structures represented by the affixes (e.g., even if we consider affixes as spell–outs of rules, it is undeniable that *–inho* provides a meaning of evaluation to the base in a word such as *cãozinho* 'small/cute dog'. The presence of the suffix actualizes the 'evaluative' meaning.). A low type frequency of the affixal chain leads to a low expectedness of the combination of the semantics provided by those affixes, which leads to a lack of semantic transparency.
3. A total semantic match is higher when affixation is pleonastic, because expectedness creates semantic transparency. Expectedness at the level of combination between affixes has an effect on expectedness at the level of combination between conceptual structures.
4. Partial phonological match is higher when affix combination is unexpected.
5. Partial phonological match is higher when affix combination has a low pattern character.

5.3 Results

5.3.1 Acceptability judgment task

As shown in Figure 1, in the acceptability judgment task acceptance rates were 97% for neutral words; 97% of frequent evaluative words; 38% for *–bil/–al–izabilidade* words; 29% of non–frequent or created evaluative words and 24% of pseudo–words.

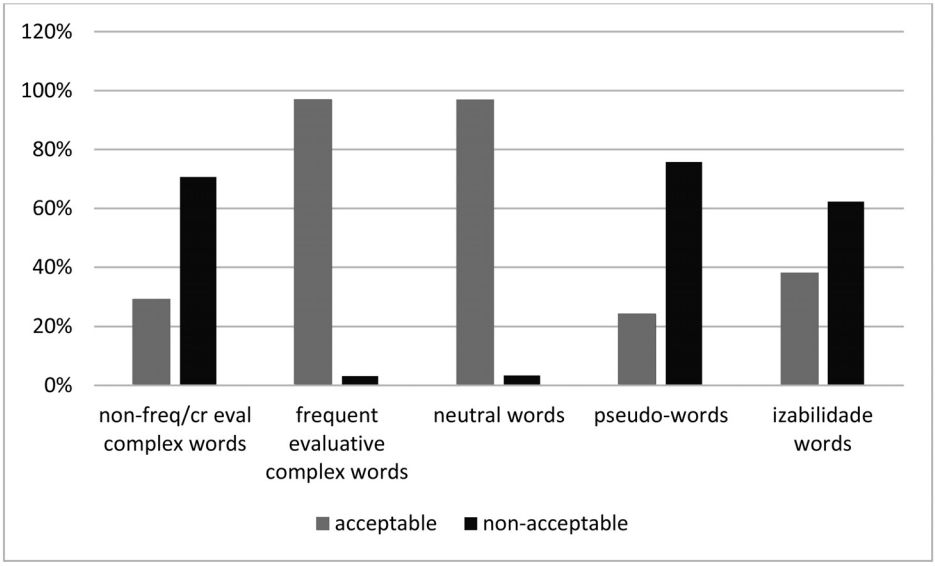

Figure 1: Results of acceptability judgment task

5.3.2 Recall task

Figure 2 summarizes results for the recall task. Neutral words obtained 95% of total lexical matches, followed by frequent evaluative words with 81%, non–frequent evaluative words (63%), *–bil/–al–izabilidade* words (30%) and pseudo–words (0%). Partial phonological mutation was higher in *–bil/–al–izabilidade* words (28%), followed by pseudo–words (25%), non–frequent or created evaluative words (10%), frequent evaluative words (5%) and neutral words (0%). Total semantic match was higher in non–frequent–created evaluative words (20%), followed by frequent evaluative words (7%) and neutral words (1%). Pseudo–words, *–bil/–al–izabilidade* words did not obtain any result of this type. Partial semantic match was higher in *–bil/–al–izabilidade* words (35%), followed by non–frequent or created evaluative words (3%), frequent evaluative words (2%), neutral words (1%) and pseudo–words (0%). Lexical category mismatch appeared in 1% of non–frequent or created evaluative words and in the same percentage of *–bil/–al–izabilidade* words. It did not appear in the other classes of words. Other real words were produced in 1% of frequent evaluative words, of neutral nouns and of *–bil/–al–izabilidade* words. No answers were obtained in 75% of pseudo–words, 5% of the *–bil/–al–izabilidade* words, and 3% of the non–frequent or created evaluative words, of the frequent evaluative words, and of neutral words.

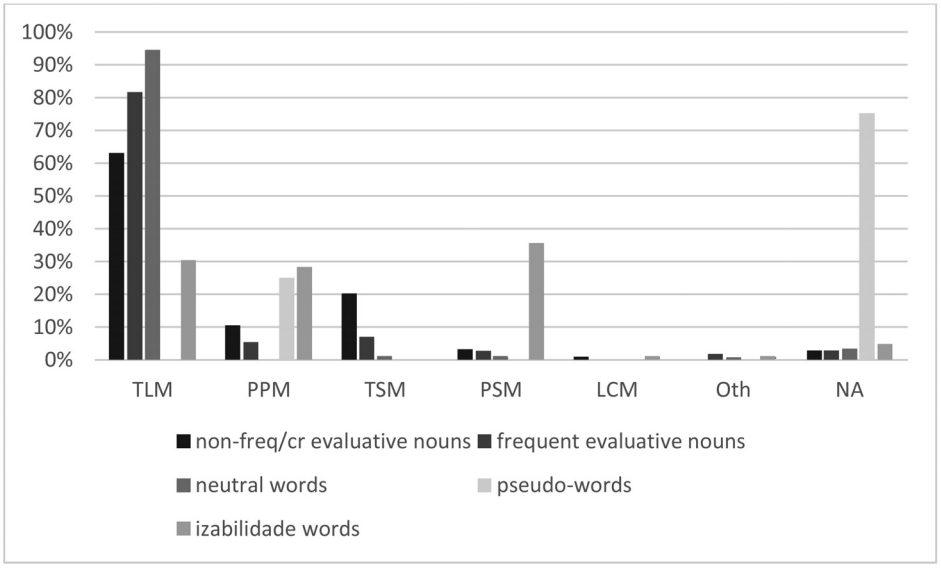

Figure 2: Results of recall task

6. General discussion

Results of the acceptability judgment task indicate that, at the level of reception, a combination of four heterocategorial derivational suffixes and a chain of pleonastic affixation in a complex word tend not to be accepted by Portuguese native speakers. This complexity seems related to the non–frequency of the word, as in *–bil/–al–izabilidade* words and in non–frequent or created evaluative words. The fact that some neutral words were classified as non–acceptable may be interpreted as a mismatch between word awareness and word usage by speakers. Because participants had academic backgrounds, but no knowledge of linguistics, there may be a tendency to confound a 'correct' usage of the language with 'acceptance'. This explanation is supported by the results obtained in relation to slang words, included in the group of neutral words, such as *bué* 'very' and *nicles* 'nothing'. These slang words led to 'non–acceptable' answers in neutral words. This may indicate that even though speakers had no linguistic training, they had a clear academic vision of the 'correct' usage of language, and considered slang words as non–words. Thus, although they recognized the word, they would reject it as not belonging to a standard discourse.

Results from the acceptability judgment task are more interesting if interpreted together with the results of the recall task. These results need to be compared across categories of words. First, we compare non–frequent or created evaluative words with frequent evaluative words. Second, we compare non–frequent or created evaluative words with *–bil/–al–izabilidade* words. A

third comparison, between pseudo–words and non–frequent/created evaluative words and *–bil/–al–izabilidade* words, is also important for our understanding. In each of the two groups, we are dealing with forms that were classified by Portuguese native speakers as non–acceptable.

6.1 Comparison between non–frequent/created evaluative words with frequent evaluative words

Non–frequent/created evaluative words and frequent evaluative words are characterized by evaluative affixation. Non–frequent/created evaluative words present pleonastic affixation, that is, the occurrence of suffixes operating in the same pattern, specifically, the evaluative–diminutive pattern. Differences between the results arise from the frequency/non–frequency of the kind of word.

1. Total lexical match is higher in frequent evaluative words than in non–frequent/created ones, which reveals a relation between frequency and expectedness. This is in accordance with prediction 1.
2. Partial phonological match is higher in non–frequent evaluative words than in frequent ones. Partial phonological matches found in the recall of these words represent productions where one or more phonological segments or syllables of the word:
   i) was/were omitted (non–frequent created word: *cadelizinha* [kɐ.dɨ.li.ˈzi.ɲɐ] instead of the stimulus *cadelinhazinha* [kɐ.dɨ.li.ɲɐ.ˈzi.ɲɐ] ‘very small she–dog’);
   ii) changed place inside the syllable or changed syllable (onset of antepenultimate syllable changed place with the onset of the penultimate syllable (*aranhizocinho* [ɐ.ɾɐ.ɲi.zu.ˈsi.ɲu] instead of the stimulus *aranhiçozinho* [ɐ.ɾɐ.ɲi.su.ˈzi.ɲu] ‘very small spider’));
   iii) was/were substituted by another one (*pequeninocinho* [pikɨninuˈsiɲu] instead of the stimulus *pequeninozinho* [pikɨninuˈziɲu] ‘very very small’).
   This difference is because the affix combination in non–frequent words is less recalled as a whole set than in frequent ones, being less expected by the speaker. This reveals that prediction 4 is correct. Semantic consequences occurred in none of these situations.
3. Total semantic match is higher in non–frequent evaluative words than in frequent ones. Non–frequent evaluative words were substituted by frequent words with the same meaning as the stimulus word. This is because the actual word was less recalled as a whole in non–frequent nouns than in frequent ones (*cãozinho* ‘small dog’ instead of *cãozinhinho*, *cadelinha*/*cadelazinha* ‘small she–dog’ instead of *cadelinhazinha*; *mesinhazita* ‘small valueless table’ instead of *mesinhitazinha*). These results indicate that the final goal of processing, which consists of semantic interpretation of the stimuli, is achieved, which is shown by the total semantic match, despite the non–identification of every morphological constituent of the word.

4. Difference in partial semantic match is insignificant. For every word, the produced word matched the semantics of 'diminutive'. Where the base of the word was substituted, the new one shared semantic features with the stimulus (e.g. *pequenino* 'very small' instead of *pequenininhozinho* 'very, very small'). This is because pleonastic affixation, such as the evaluative one, leads to semantic expectedness in the hearer that enables him/her to identify the meaning of the word, even in relation to non–frequent words, which is in accordance with prediction 2.
5. Other real words were produced in the substitution of frequent words. (*passinho* (*passo–inho*, where *passo* means 'step'), instead of the stimulus *passarinho* (*pássaro–inho*, where *pássaro* refers to 'bird'). In the cases where this occurred, a phonological misunderstanding of the base seems to have occurred. No result of this kind was found concerning non–frequent words.
6. There was a low percentage of 'no answers' for both groups of words, equivalent across the groups. The same was observed for neutral words. This is because even in non–frequent words, the evaluative affixation was recognized, which permitted an identification of the word, even with changes to its structures. Pleonastic affixation leads to expectable semantics carried by the affixes, which enhances the semantic transparency of the word.

6.2 Comparison between non–frequent or created evaluative words with –bil/–al–izabilidade words

Both groups represent non–frequent words. The difference between them is that non–frequent or created evaluative words have pleonastic suffixation, whilst *–bil/–al–izabilidade* words have heterogeneous suffixation.

– **Total lexical match** is higher in non–frequent/created evaluative words than in *–bil/–al–izabilidade* words. This is because it is easier to recall a complex word with pleonastic suffixation, where each one of the derivational suffixes actualizes the same kind of meaning, and the same kind of derivational pattern, than a complex word that contains several suffixes each one actualizing a different meaning and a different pattern, even if both groups of words are non–frequent ones. The presence of pleonastic affixation reveals a prediction effect, because the meaning of each of the identified morphemes enables the prediction of the same meaning in the other morphemes, as long as they actualize the same pattern. Pleonastic affixation reinforces expectedness and semantic transparency inside the word. Because of this internal invocation of semantic structures that has an effect on expectedness, the effect of the non–frequent character of the words is minimalized when compared with other non–frequent words that lack pleonastic affixation. This indicates that Prediction 2 is correct.

Due to the complexity of processing a non–frequent chain that contains four suffixes lacking a pleonastic effect, recalling words with a total lexical

match would have required a memorized affix combination or the recognition of a word–formation pattern in that combination.

– **Partial phonological mutation** is higher in *–bil/–al–izabilidade* words than in non–frequent/created evaluative words. Again, this is owing to the difficulty in recognizing the whole word and, at the same time, in *–bil/–al–izabilidade* words, to the difficulty in recognizing the derivational morphemes involved in a complex word with no pleonastic suffixation, if the affixal combination is not an expected one, which decreases affixal salience.

All the situations of partial phonological mutation in *–bil/–al–izabilidade* words indicate a difficulty in recognizing the morphological constituents as identities. This is due to the unexpectedness of the combination of those morphemes, because the combination is not frequent and lacks pleonastic affixation. This shows that affixal salience (Laudanna and Burani 1995) also depends on the expectedness of the affixal combination. This is in accordance with predictions 4 and 5.

Partial phonological mutation occurs as:

a) Substitution of morphological components by others that do not operate in the same word–formation pattern (*famarialização* instead of *familiarizabilidade*, where the produced form contains the suffix *–ção* that operates in the formation of event deverbal nouns. Notice that *famarialização* does not correspond to any word in Portuguese, nor does the segment at the position of the base correspond to any base of Portuguese. Bearing this in mind, it is not possible to consider that there was a semantic mutation);

b) Omission of phonological segments that do not coincide with morphological frontiers (in *–bil/–al–izabilidade* words: *familiaribilidade* instead of *familiari**za**bilidade*; *comerçalizabidade*/ *comercizibidade* instead of *comerc**i**alizabi**li**dade*; *contabilizalidade* instead of *contabiliza**bi**lidade*). Notice that the omitted phonological segments do not coincide with morphological frontiers (*famili–ar–i**z**–**a**–bil–idad–e*, *comerc**i**–al–iz–a–bi**l**–**i**dad–e*; *conta–bil–iz–a–**bi**l–idad–e*), but with syllables (*fa.mi.li.a.ri.**za.**bi.li.da.de*; *co.mer.ci.a.li.za.bi.**li**.da.de*; *con.ta.bi.li.za.**bi.**li.da.de*);

c) Change of place of a segment inside a syllable or between syllables (in *–bil/–al–izabilidade* words: *famili**razi**dade* instead of *familiar**iza**bilidade*);

d) Omission of segments that correspond to morphological units (*familirazidade* instead of *familiariza**bil**idade*; *comercializidade* instead of *comerci–al–iz–**a–bil**–idad–e*; *legalizidade* instead of *legal–i–z–**a–bil**–idad–e*);

e) Insertion of phonological segments that do not appear at the stimulus (*comerç**acio**bilidade* instead of *comercializabilidade*);

f) Substitution of phonological segments by others that are phonologically equal to the ones existing in the surrounding syllables (*familiariz**i**bilidade* instead of *familiar**iza**bilidade*);

As can be seen by the examples, the different cases may occur together in one production.

– **Total semantic match** is high in non–frequent/created evaluative words and does not occur in *–bil/–al–izabilidade* words, which confirms predic-

tion 3. Even when the word is not totally recalled, the existence of pleonastic suffixation permits changes of suffixes, which, being synonymous and operating in the same word formation pattern, allow for the preservation of the evaluative meaning of the word (*mesinhazita* 'small valueless table'; *mesinhitazinhinha* 'very very small valueless table' instead of *mesinhitazinha* 'very small valueless table'). This does not occur with *–bil/–al–izabilidade* words, where a change in the morphological segments causes differences in the meaning of the words, as the results are classified as partial semantic matches (e.g. *sociabilização* 'event of turning someone/becoming sociable' instead of *socializabilidade* 'quality of being able to socialize'). For the same reasons, partial change in syllables may be semantically indifferent in evaluative words, because these have pleonastic affixation. This is not so in heterocategorial suffixation, where a change in the syllable has semantic consequences or leads to the production of an impossible word (**contabilizalidade* instead of *contabilizabilidade*; **materiazabilidade* instead of *materializabilidade*).

– **Partial semantic match.** The distance between pleonastic suffixation and complex heterocategorial suffixation in terms of semantic effects has consequences in the results obtained concerning partial semantic match. In *–bil/–al–izabilidade* words, partial semantic match is higher (35%) than in non–frequent/created evaluative words (3%). The results in the last group of words should be interpreted in relation to the results obtained in total semantic matches (20%). Again, in pleonastic suffixation, changes in the produced forms that result in real/possible words maintain the evaluative meaning of the word, whilst in heterocategorial formations changes in the produced forms that result in real/possible words tend to alter the meaning conveyed by the stimulus.

In fact, in *–bil/–al–izabilidade* words, the stimuli were substituted by other real or possible words of Portuguese with a semantic similarity with the stimulus (*familiaridade* 'familiarity' instead of *familiarizabilidade* 'quality of being able to turn something familiar'; *contratualização* 'event of settle something/someone by a contract', *contratuabilidade* 'quality of being able to be settled by a contract', *contratualidade* 'quality of being contractual' instead of *contratualizabilidade* 'quality of being able to make something by a contract'; *comerciabilidade* 'quality of being able to be traded' instead of *comercializabilidade* 'quality of being able to be commercialized'; *globalização* 'globalization' instead of *globalizabilidade* 'quality of being able to be globalized'; *contabilidade* 'accountancy, accounting', *contabilização* 'accounting' instead of *contabilizabilidade* 'quality of being able to be accounted'; *materialidade* 'materiality' instead of *materializabilidade* 'quality of being able to be materialized'; *sociabilidade* 'sociability', *sociabilização* 'event of turning someone sociable' instead of *socializabilidade* 'quality of being able to be socialized').

Two relevant situations are observable in the examples quoted above:

a) Preservation of the ending suffix *–idade* and concomitant loss of the suffix *–iz–*; or

b) Preservation of *–iz–* which leads to add a final suffix –ção with concomitant *–idade*.

This shows that, as *–bil/–al–izabilidade* words do not contain an expected combination of affixes, they do not correspond to realizations of a word–formation pattern. Words produced with

a) Preservation of *–idade* and concomitant loss of suffix *–iz–*;

b) Preservation of *–iz–* and addition of the suffix –ção

show us that speakers tend to follow patterns that are expected, and that patterns *–iz–ção* and *–idade* lacking *–iz–*, which exist in their minds, have a stronger pattern character than *–bil/–al–izabilidade* has. The pattern producing quality nouns with the suffix *–idade* has a strong pattern character. This is suggested by the following: Recalling the *–idade* suffix from the stimuli leads to the abandonment of *–iz–*, because this is not a frequent association in Portuguese words. A second pattern is also strong: The pattern that constructs event nouns with the suffix *–ção* from *–iz–* verbs. The bias towards this pattern is shown by the recalling of suffix *–iz–* from the stimuli leading to the pattern formation of nouns with suffix *–ção*. This suggests a prediction effect on the level of morphological structures inside the word, according to which the recognition of a morpheme enables the prediction of another morpheme, which, although being absent from the stimulus, is associated with the first morpheme in the mental lexicon. (Cf. Caballero (2010) who reports that in Choguita Rarámuri, a Uto–Aztecan language, during informant sessions, due to priming effects, some affix orderings become fixed, serving as bases for other constructions, even when those affix orderings were not expected, since they do not reflect the intended semantics.)

The results concerning Partial semantic match reveals that there is a correlation between Partial semantic matches and the low pattern character of the affixal chain. We had not anticipated that correlation in our predictions. Nevertheless, based on the results, we can state now that Partial semantic match is higher when affix combination has a low pattern character.

For both sets of words, 'no answers' did not get a significant result. This indicates that in both groups of words at least some morphological segments of the word were identified. This interpretation makes sense if we compare these results with results obtained for pseudo–words. For the pseudo–words, 75% of the answers were 'no answers', where no identification of the word or parts of the word was achieved.

6.3 Comparison between pseudo–words and non–frequent/created evaluative words and –bil/–al–izabilidade *words*

A comparison between pseudo–words and non–frequent/created evaluative words and *–bil/–al–izabilidade* words is relevant, bearing in mind that, for the three groups of forms, the 'non–acceptable' classification obtained in the acceptability judgment task yielded high percentages (pseudo–words: 76%; non–frequent/created evaluative words: 71%; *–bil/–al–izabilidade* words: 62%). In the three groups, we are dealing with forms that were generally unacceptable to Portuguese native speakers. Interestingly, in the recall task, pseudo–words were treated differently from *–bil/–al–izabilidade* words and non–frequent/cre-

ated evaluative words. Of all the groups of words, pseudo–words obtained the highest percentage of no answers (75%), against 5% for *–bil/–al–izabilidade* words and 3% for non–frequent/created evaluative words. These results are similar to those obtained for neutral words (3%).

This discrepancy among classes of words that are similarly classified as non–acceptable in the acceptability judgment task is because of the identification of some of the morphemes of the words, which does not occur with pseudo–words that were structured in a way to avoid misinterpretations of eventual morphemes. For instance, pseudo–words included in the stimuli were the forms *quirilete*, *vontoco*, *citisco*, *socir*, *pomponipom*, *livivi*, *fanti*, which do not contain phonological combinations interpretable as morphological constituents in Portuguese or as actualizing Portuguese word–formation patterns. In different terms, *–bil/–al–izabilidade* words and non–frequent/created evaluative words, although classified as non–acceptable with a high percentage in the acceptability judgment task, contain morphological constituents (either lexical or affixal constituents) that may be interpreted as actualizing derivational patterns and/or bases of Portuguese.[6]

7. Conclusions

In a study concerning compounds, Bell and Schäfer (2013) stipulated that not only the transparency of the constituents of the compound, but also the semantic relation between those constituents, contributes to the semantic transparency of the compound. Bell and Schäfer (2016: 163) test "the hypothesis that perceived transparency is correlated both with the expectedness of the constituents themselves and with the expectedness of this relation".

We extended those findings to affixed words, viz. affixed words with four derivational affixes. Our target was constituted by *–bil/–al–izabilidade* words, containing four derivational affixes, which, although appearing in Google searches, have a low frequency in Portuguese. Our main aim was to determine if the specific combination of affixes, as a whole, represented in those words constitutes a word–formation pattern in contemporary Portuguese. Constituting a word–formation pattern requires adherence to structural constraints and processing conditions. These processing conditions result from the correlation between frequency, expectedness towards the affixal combination, which increase semantic transparency (Bell and Schäfer (2016) and affixal salience (Laudanna and Burani 1995) and enable the speaker to recognize the affixal combination to analyse and form words. By emphasizing this correlation, this study contributes to understand the limits of multiple affixation.

6 Remember what we have said about total semantic matches in evaluative words and partial semantic matches in *–bil/–al–izabilidade* words. In the first set, formal changes occurred, but the semantics of evaluation was maintained. In the second, the presence of heterocategorial suffixation makes difficult a total semantic match when formal changes occur, but a partial semantic match is obtained. Specifically, the produced forms showed a tendency either in the direction of the word–formation pattern of quality nouns with the suffix *–idade*, or in the direction of *–iz–* pattern, which led to productions with the suffix *–ção*.

We conclude that *–bil/–al–izabilidade* words do not constitute a word–formation pattern and that the affixal combination has a low pattern character. In *–bil/–al–izabilidade* words the combination of affixes is not expected, which reduces semantic transparency and the affixal salience of the words containing it (Caramazza et al. (1998), Schreuder and Baayen (1995), Frauenfelder and Schreuder (1991) and Ji et al. (2011)). This kind of words contrast with non–frequent pleonastic evaluative words. Albeit non–frequent, pleonastic evaluative words contrast with *–bil/–al–izabilidade* words in their semantic transparency.

Some of the results may be interpreted by considering a prediction effect at a syntagmatic level. In this case, the prediction effect takes place inside a particular autonomous lexical unit (a complex word), where the presence of a morpheme activates the recognition of another morpheme, which is expected to occur (cf. Caballero 2010). This expectedness is related to combinations of affixes that constitute word–formation patterns and, thus, works on the level of morphological structures, as shown by the results of a recall task of *–bil/–al–izabilidade* words. Patterns such as the combination *–iz–ção* and quality noun formation in *–idade* (which perform the activation of the absence of suffix *–iz–*) – observed in the recall task results – show a stronger pattern character than the combination *–bil/–al–iz(a)–bil–idade*, which leads us to consider that the low frequency and acceptance of the latter type of word is associated with its non–pattern character. Although the frequency of the affix chain does not determine the easiness of the identification of morphemes, since the higher the frequency of the chain the more the combination will be observed a whole, the point is that, in *–bil/–al–izabilidade* words, the identification of the morphemes is not favoured due to the failure to recognize the specific combination of affixes as a pattern. That failure motivates a recategorization of units according to expected patterns (*–iz–ção* or *–idade* without *–iz–*).

Another prediction effect is suggested in pleonastic affixation. Here, expectedness works at the level of semantic structures and not at the level of morphology. The presence of an evaluative suffix may turn predictable evaluative semantic structures expressed by the surrounding suffixes. This emphasises studies such as Ji et al. (2011: 419), who highlight the importance of semantics to the processing of words. It also stresses works such as Gagné and Spalding (2004, 2009, 2010), who propose that the reconstruction of meaning depends on data of the meanings of constituents, when available.

The fact that *–bil/–al–izabilidade* words are not instances of a word–formation pattern is a consequence of low frequency, low expectedness, lack of semantic transparency and affixal salience (Laudanna and Burani 1995). However, as Bell and Schäfer (2016: 195, footnote 10) suggest, "There could be also a positive feedback effect, whereby users are more likely to choose to use items they themselves would find easier to process, so that ease of processing conversely leads to greater frequency of use". Thus, because the combination does not constitute a pattern, it is seldom used by speakers. One of the anonymous reviewers pointed out that if the words under study do exist, if they are produced and interpreted, this should be taken as «witness of the psychological reality of a corresponding word formation pattern». However, we believe

that the existence of a constructed word does not imply the existence of a corresponding pattern, as compound words show. Compound words may convey an idiosyncratic meaning, which does not instantiate a pattern, but they are still interpretable and produced.

All of these factors have consequences on the processing of these words. Despite the adherence of *–bil/–al–izabilidade* words to structural constraints, processing of these words encounters difficulties. In a parallel dual–route model (Baayen, Dijkstra and Schreuder 1997; Schreuder and Baayen 1997), the decomposed route would be favoured owing to the combination of two factors: the very low–frequency of the word and its morphological complexity degree (Niswander–Klement and Pollatsek 2006). However, there are obstacles in this route related to the low semantic transparency and low affixal salience of the affixal combination, which stem from the extension of the affixal combination. Considering that the goal of a speaker's morphological analysis is to create meaning (Libben 2015), that morphological structure serves the computation of meaning (Schreuder and Baayen 1995: 132), and that the presence of morphemes in a word expresses a conceptual category, the quantity of morphological segments with neither referential nor pleonastic semantics makes morphological parsing difficult. There is another obstacle to whole–word processing: Because this combination of suffixes has a very low frequency, it does not correspond to predictable/expected combinations in the speaker's mind (Hawkins and Blakeslee 2004; Plag and Baayen 2009).

One type of multiple affixation contrasts with those cases:

Multiple affixation in evaluative words. Pleonastic evaluative words are easily processed, because: 1) evaluative affixes have higher semantic transparency than suffixes such as *–al–*, *–bil–*; and 2) in pleonastic evaluative formations each one of the suffixes repeats the semantic information of its precedent. In this case, multiple affixation highlights information needed to process the word, and the expectedness towards the constituents of the word is reinforced by semantic structures.

Although the experiments we carried out are not sufficient for making all–encompassing assertions, we may tentatively infer from our results and predict that the affixal combination *–bil–/al–iz(a)–bil–idade* has a low pattern character. Because this combination has a low frequency, it is not often experienced by the speaker, and thus there is limited experience of this combination to constitute a pattern in the speaker's mind (Hawkins and Blakeslee 2004; Plag and Baayen 2009). Conversely, as this combination does not constitute a pattern, it is difficult to recognize words containing that combination using mental data that would correspond to the pattern. This is in accordance with Tomasello (2000), who stresses the role of specific linguistic phenomena represented in the speaker's mind in the construction of abstract schemata and with Baayen (2007: 84) in his idea that the processing of complex words benefits from the brain "keep[ing] track of detailed combinatorial probabilities".

We suggest that future experiments are needed to demonstrate the degree to which expectedness of an affixal chain determines its constitution as a word–formation pattern and its productivity.

Our study is in accordance with multiple–route models, such as the one proposed by Kuperman et al. (2010), in the sense that our data brings evidence to the assumptions that "morphological structure [functions] as a conglomerate of sources of information, which contribute – to a different extent – to the recognition of polymorphemic words," and that "morphemes, combinations of morphemes, and morphological paradigms and structurally complex words [function] as sources of morphological information" (Kuperman et al. 2010: 94).

References

Allen, M. (1978). *Morphological investigations*. PhD. Dissertation. University of Connecticut: Storrs.

Andrews, S. (1986). Morphological influences on lexical access: Lexical or nonlexical effects? *Journal of Memory and Language*, 25, 726–740. https://doi.org/10.1016/0749-596X(86)90046-X

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology*, 16, 285–311. https://doi.org/10.1080/09541440340000123

Aronoff, M. (1976). *Word–formation in generative grammar*. Cambridge, MA: The MIT Press.

Baayen, H. (1992). Quantitative aspects of morphological productivity. In G. Booij and Jan van Marle (Eds.), *Yearbook of Morphology 1991*. Dordrecht: Kluwer, 109–149.

Baayen, H. (1993). On frequency, transparency and productivity. In Geert Booij & J. van Marle (Eds.), *Yearbook of Morphology 1992* (pp. 181–208). Dordrecht: Kluwer.

Baayen, H. (2007). Storage and processing in the mental lexicon. In G. Jarema & G. Libben (Eds.), *The mental lexicon. Core perspectives* (pp.81–104). Amsterdam: Elsevier.

Baayen, H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: evidence for a parallel dual–route model. *Journal of Memory and Language*, 37, 94–117. https://doi.org/10.1006/jmla.1997.2509

Baayen, H., & Schreuder, R. (2000). Towards a psycholinguistic computational modeling of morphological parsing. *Philosophical Transactions of the Royal Society: Mathematical, Physical and Engineering Sciences. Computers, Language and Speech: Formal Theories and Statistical Data*, 358(1769), 1281–1293.

Baayen, H., Wum, L. H., & Aycock, J. (2007). Lexical dynamics for low–frequency complex words. A regression study across tasks and modalities. *The mental lexicon*, 2(3), 419–463. https://doi.org/10.1075/ml.2.3.06baa

Baeskow, H. (2015). Rules, patterns and schemata in word–formation. In P. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), *Word–formation. An international handbook of the languages of Europe.* Vol. 2 (pp. 803–821). Berlin: Mouton de Gruyter.

Barz, I. (2016). German. In P. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), *Word–formation. An international handbook of the languages of Europe.* Vol. 4 (pp. 2387–2411). Berlin: Mouton de Gruyter.

Bauer, L. (1983). *English word–formation*. Cambridge: Cambridge University Press.

Bauer, L. (2001). *Morphological productivity*. Cambridge: Cambridge University Press.

Bauer, L. (2005). Productivity: Theories. In P. Štekauer, & R. Lieber (Eds.), *Handbook of word–formation* (pp. 315–334). Dordrecht: Springer.

Bauer, L., Lieber, R. & Plag, I. (2013). *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.

Bauer, L., & Valera, S. (Eds.) (2005). *Approaches to conversion/zero–derivation*. Munich: Waxmann.

Beard, R. (1995). *Lexeme–morpheme base morphology: a general theory of inflection and word formation*. New York: State University of New York Press.

Beauvillain, C. (1996.) The integration of morphological and whole–word information during eye–fixations on prefixed and suffixed words. *Journal of Memory and Language*, 35, 801–820. https://doi.org/10.1006/jmla.1996.0041

Bell, M. J., & Schäfer, M. (2013). Semantic transparency: challenges for distributional semantics. In A. Herbelot, R. Zamparelli, & G. Boleda (Eds.), *Proceedings of the IWCS 2013 workshop: Towards a formal distributional semantics. Potsdam: Association for Computational Linguistics* (pp. 1–10). http://www.aclweb.org/anthology/W13–0601. Accessed 23 February 2016.

Bell, M. J., & Schäfer, M. (2016). Modelling semantic transparency. *Morphology* http://link.springer.com/article/10.1007/s11525–016–9286–3. Accessed 15 May 2016. https://doi.org/10.1007/s11525-016-9286-3

Bertram, R., Schreuder, R., & H. Baayen (2000). The balance of storage and computation in morphological processing: the role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 489–511. https://doi.org/10.1037/0278-7393.26.2.489

Blevins, J. (2016). *Word and paradigm morphology*. Oxford: Oxford University Press.

Bloomfield, L. (1933). *Language*. New York: Henry Holt.

Booij, G. (2010). *Construction morphology*. Oxford: Oxford University Press.

Booij, G. 2016. Dutch. In P. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), *Word–formation. An international handbook of the languages of Europe.* Vol. 4 (pp. 2427–2451). Berlin: Mouton de Gruyter.

Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language and Cognitive Processes*, 2, 217–227. https://doi.org/10.1080/01690968708406932

Burani, C., & Thornton, A. (2003). The interplay of root, suffix and whole–word frequency in processing derived words. In R. H. Baayen & R. Schreuder (Eds.), *Morphological structure in language processing* (pp. 157–208). Berlin: Mouton de Gruyter.

Burzio, L. (1991 [1995]). On the metrical unity of Latinate affixes. In H. Campos & P.M. Kempchinsky (Eds.), *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero* (pp. 1–24). Washington: Georgetown University Press.

Burzio, L. (2002). Surface–to–surface morphology. When your representations turn into constraints. In P. Boucher (Ed.), *Many morphologies* (pp. 34–44). Somerville; Mass.: Cascadilla.

Butterworth, B. (1993). Lexical representation. In B. Butterwoth (Ed.), *Language production*. Vol. 2 (pp. 257–294). San Diego, CA: Academic Press.

Caballero, G. (2010). Scope, phonology and morphology in an agglutinating language: Choguita Rarámuri (Tarahumara) variable suffix ordering. *Morphology*, 20, 165–204. https://doi.org/10.1007/s11525-010-9147-4

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition* 28, 297–332. https://doi.org/10.1016/0010-0277(88)90017-0

Carstairs–McCarthy, A. (1992). *Current morphology*. London and New York: Routledge.

Corbin, D. (1987). *Morphologie dérivationnelle et structuration du lexique*. 2 vols. Tübingen: Niemeyer.

Diependaele, K, Sandra, D., & Grainger, J. (2005). Masked cross–modal morphological priming: unravelling morpho–ortographic and morpho–semantic influences in early word recognition. *Language and Cognitive Processes*, 20, 75–114. https://doi.org/10.1080/01690960444000197

Dressler, W. U., Dziubalska–Kołaczyk, K., Gagarina, N., & Kilani–Schoch, M. (2015). Reduplication, repetition, hypercharacterization, and other affix–doubling in child–language. In S. Manova (Ed.), *Affix ordering across languages and frameworks* (pp. 259–275). Oxford: Oxford University Press.

Frauenfelder, U. H., & Schreuder, R. (1991). Constraining psycholinguistic models of morphological processing and representation: the role of productivity. In G. Booij, & J. van Marle (Eds.), *Yearbook of morphology 1990* (pp. 165–138). Dordrecht: Kluwer.

Frisson, S., Niswander–Klement, E., & Pollatsek, A. (2008). The role of semantic transparency in the processing of English compound words. *British Journal of Psychology*, 99, 87–107. https://doi.org/10.1348/000712607X181304

Gaeta, L. (2005). Combinazioni di suffissi in italiano. *La formazione delle parole*. In M. Grossmann, & A. M. Thornton (Eds.), *Atti del XXXVII Congresso Internazionale di Studi della Società di Linguistica Italiana* (L'Aquila, 25–27/9/2003) (pp. 229–247). Roma: Bulzoni.

Gaeta, L., & Ricca, D. (2015). Productivity. In P. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), *Word–formation. An international handbook of the languages of Europe*. Vol. 2 (pp. 842–858). Berlin: Mouton de Gruyter.

Gagné, C., & Spalding, T. L. (2004). Effect of relation availability on the interpretation and access of familiar noun–noun compounds. *Brain and Language*, 90(1–3), 478–486. https://doi.org/10.1016/S0093-934X(03)00459-0

Gagné, C., & Spalding, T. L. (2009). Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language*, 60, 20–35. https://doi.org/10.1016/j.jml.2008.07.003

Gagné, C., & Spalding, T. L. (2010). Relational competition during compound interpretation. In S. Scalise, & I. Vogel (Eds.), *Cross–disciplinary issues in compounding* (pp. 287–300). Amsterdam: John Benjamins.

Gardani, F. (2015). Affix pleonasm. In P. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), *Word–formation. An international handbook of the languages of Europe*. Vol. 1 (pp.537–550). Berlin: Mouton de Gruyter.

Giegerich, H. (1999). *Lexical strata in English*. Cambridge: Cambridge University Press.

Giraudo, H., & Grainger, J. (2001). Priming complex words: evidence for supralexical representation of morphology. *Psychonomic Bulletin & Review*, 8, 127–131. https://doi.org/10.3758/BF03196148

Halle, M. (1973). Prolegomena to a theory of word formation. *Linguistic Inquiry*, 4, 3–16.

Halle, M., & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. L. Hale, S. J. Keyser, & S. Bromberger (Eds.), *The view from building 20: Essays in Linguistics in honour of Sylvain Bromberger* (pp. 111–176). Cambridge: Cambridge University Press.

Hawkins, J., & Blakeslee, S. (2004). *On intelligence*. New York: Henry Holt and Company.

Hay, J. B., & Baayen, H. (2002). Parsing and productivity. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology 2001* (pp. 203–235). Dordrecht: Kluwer Academic Publishers.

Hay, J. B., & Baayen, H. (2005). Shifting paradigms. Gradient structure in morphology. *Trends in Cognitive Sciences*, 9, 342–348. https://doi.org/10.1016/j.tics.2005.04.002

Hay, J. B., & Plag, I. (2004). What constrains possible suffix combinations? On the interaction of grammatical and processing restrictions in derivational morphology. *Natural language and linguistic theory*, 22, 565–596. https://doi.org/10.1023/B:NALA.0000027679.63308.89

Ji, H., Gagné, C., & Spalding, T. (2011). Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds. *Journal of Memory and Language*, 65, 406–430. https://doi.org/10.1016/j.jml.2011.07.003

Kumar, A., Mittal, V., & Kulkarni, A. (2010). Sanskrit Compound Processor. In G. N. Jha (Ed.), *Sanskrit Computational Linguistics: 4th International Symposium* (New Delhi, 10–12/12/2010): *Proceedings (Volume 6465 of Lecture Notes in Computer Science/Lecture Notes in Artificial Intelligence Series* (pp. 111–176). Springer.

Kuperman, V. Bertram, R., & Baayen, H. (2008). Morphological dynamics in compound processing. *Language and Cognitive Processes*, 23, 1089–1132. https://doi.org/10.1080/01690960802193688

Kuperman, V., Bertram, R., & Baayen, H. (2010). Processing trade–offs in the reading of Dutch derived words. *Journal of Memory and Language*, 62, 83–97. https://doi.org/10.1016/j.jml.2009.10.001

Kuperman, V., Schrueder, R., Bertram, R., & Baayen, H. (2009). Reading polymorphemic Dutch compounds: toward a multiple route of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 876–895. https://doi.org/10.1037/a0013484

Laudanna, A., & Burani, C. (1995). Distributional properties of derivational affixes: Implications for processing. In L.B. Feldman (Ed.), *Morphological Aspects of Language Processing: Cross–Linguistic Perspectives* (pp. 345–364). Hillsdale: Lawrence Erlbaum Associates.

Laudanna, A., Burani, C., & Cermele, A. (1994). Prefixes as processing units. *Language and Cognitive Processes*, 9(3), 295–316. https://doi.org/10.1080/01690969408402121

Lehtonen, M., Cunillera, T., Rodriguez–Fornells, A., Hulten, A., Tuomainen, J., & Laine, M. (2007). Recognition of morphologically complex words in Finnish: evidence from event–related potentials. *Brain Research*, 1148, 123–137. https://doi.org/10.1016/j.brainres.2007.02.026

Libben, G. (1994). How is morphological decomposition achieved? *Language and Cognitive Processes*, 9(3), 369–391. https://doi.org/10.1080/01690969408402124

Libben, G. (1998). Semantic transparency in the processing of compounds: consequences for representation, processing and impairment. *Brain and Language* 61, 30–44. https://doi.org/10.1006/brln.1997.1876

Libben, G. (2006). Why study compound processing: an overview of the issues. In G. Libben, & G. Jarema (Eds.), *The representation and processing of compound words* (pp. 1–22). Oxford University Press.

Libben, G. (2010). Compound words, semantic transparency, and morphological transcendence. *Linguistiche Berichte; Sonderheft*, 17, 317–330.

Libben, G. (2015). Word–formation in psycholinguistics and neurocognitive research. In P. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), *Word–formation. An international handbook of the languages of Europe*. Vol 1. (pp. 203–217). Berlin: Mouton de Gruyter.

Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. (2003). Compound fracture: The role of semantic transparency and morphological headedness. *Brain and Language*, 84, 50–64. https://doi.org/10.1016/S0093-934X(02)00520-5

Lieber, R. (1981). *On the organization of the lexicon*. Bloomington: Indiana University Linguistic Club.

Manova, S. (2010a). Suffix combinations in Bulgarian: Parsability and hierarchy–based ordering. *Morphology*, 20(1), 267–296. https://doi.org/10.1007/s11525-010-9148-3

Manova, S. (2010b). Introduction. In S. Manova (Ed.), *Affix ordering across languages and frameworks* (pp. 1–14). Oxford: Oxford University Press.

Manova, S. (Ed.) (2011). *Affixes and Bases. Word Structure*, 4(2). Thematic Issue.

Manova, S., & Aronoff, M. (Eds.). (2010a). *Affix Combinations. Morphology*, 20(1). Special Issue.

Manova, S., & Aronoff, M. (2010b). Modeling affix order. *Morphology*, 20(1), 109–131. https://doi.org/10.1007/s11525-010-9153-6

Marslen–Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, 101(1), 3–33. https://doi.org/10.1037/0033-295X.101.1.3

McQueen, J., & Cutler, A. (1998). Morphology in word recognition. In A. Spencer, & A. Zwicky (Eds.), *The handbook of morphology* (pp. 406–427). Oxford: Blackwell.

Niswander–Klement, E., Pollatsek, A. (2006). The effects of root frequency, word frequency, and length on the processing of prefixed English words during reading. *Memory and Cognition*, 34(3), 685–702. https://doi.org/10.3758/BF03193588

Niswander, E., Pollatsek, A., & Rayner, K. (2000). The processing of derived and inflected words during reading. *Language and Cognitive Processes*, 15, 389–420. https://doi.org/10.1080/01690960050119643

Pereira, R. (2008). *Formação de verbos em português. Formação heterocategorial*. München: Lincom.

Pereira, R. (2016). Formação de verbos. In G. Rio–Torto, A. S. Rodrigues, S. Ribeiro, R. Pereira, & I. Pereira, *Gramática derivacional do português* (pp. 297–355). Coimbra: Imprensa da Universidade de Coimbra, 2nd ed.

Plag, I. (1999). *Morphological productivity. Structural constraints in English derivation*. Berlin and New York: Mouton de Gruyter.

Plag, I. (2003). *Word–formation in English*. Cambridge: Cambridge University Press.

Plag, I., & Baayen, H. (2009). Suffix ordering and morphological processing. *Language*, 85(1), 109–152. https://doi.org/10.1353/lan.0.0087

Pollatsek, A., Slattery, T., & Juhasz, B. J. (2008). The processing of novel and lexicalized prefixed words in reading. *Language and Cognitive Processes*, 23, 1133–1158. https://doi.org/10.1080/01690960801945484

Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: The MIT Press.

Rainer, F. (2005). Constraints on productivity. In P. Štekauer, & R. Lieber (Eds.), *Handbook of word–formation* (pp. 335–352). Dordrecht: Springer.

Rainer, F. (2015). Mechanisms and motives of change in word–formation. In P. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), *Word–formation. An international handbook of the languages of Europe*. Vol. 2 (pp. 1761–1781). Berlin: Mouton de Gruyter.

Rio–Torto, G. (2016). Nomes deadjetivais. In G. Rio–Torto, A. S. Rodrigues, S. Ribeiro, R. Pereira, & I. Pereira, *Gramática derivacional do Português* (pp. 136–173). Coimbra: Imprensa da Universidade de Coimbra, 2nd ed.

Rio–Torto, G., Rodrigues, A. S., Ribeiro, S., Pereira, R. & Pereira, I. (2016). *Gramática derivacional do português*. Coimbra: Imprensa da Universidade de Coimbra.

Rodrigues, A. S. (2001). *A construção de postverbais em português*. Porto: Granito Editores e Livreiros.

Rodrigues, A. S. (2009). Portuguese converted deverbal nouns: constraints on their bases. *Word Structure*, 2, 69–107. https://doi.org/10.3366/E1750124509000300

Rodrigues, A. S. (2013). Is conversion a lexical or a syntactic process of word formation? *Linguística: Revista de Estudos Linguísticos da Universidade do Porto*, 8, 89–120.

Rodrigues, A. S. (2014). Causative eventive chains and selection of affixes in Portuguese nominalisations. *Lingue e Linguaggio*, XIII(1), 159–184. https://doi.org/10.1418/78406

Rodrigues, A. S. (2015). *A gramática do léxico: morfologia derivacional e o léxico mental*. München: Lincom.

Rodrigues, A. S. (2016a) Noções basilares sobre a morfologia e o léxico. In G. Rio–Torto, A. S. Rodrigues, S. Ribeiro, R. Pereira, & I. Pereira, *Gramática derivacional do Português* (pp. 35–133). Coimbra: Imprensa da Universidade de Coimbra, 2nd ed.

Rodrigues, A. S. (2016b). Nomes deverbais. In: G. Rio–Torto, A. S. Rodrigues, S. Ribeiro, R. Pereira, & I. Pereira, *Gramática derivacional do português* (pp. 176–240). Coimbra: Imprensa da Universidade de Coimbra, 2nd ed.

Rodrigues, A. S. (2016c). Adjetivos deverbais. In: G. Rio–Torto, A. S. Rodrigues, S. Ribeiro, R. Pereira, & I. Pereira, *Gramática derivacional do português* (pp. 261–296). Coimbra: Imprensa da Universidade de Coimbra, 2nd ed.

Ryan, K. (2010). Variable affix ordering: grammar and learning. *Language*, 86(4), 758–791.

Saarinen, P., & Hay, J. (2014). Affix ordering in derivation. In R. Lieber, & P. Štekauer (Eds.), *The Oxford handbook of derivational morphology* (pp. 370–383). Oxford: Oxford University Press.

Scherer, C. (2015). Change in productivity. In P. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), *Word–formation. An international handbook of the languages of Europe*. Vol. 2 (pp. 1781–1793). Berlin: Mouton de Gruyter.

Schreuder, R., & Baayen, H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 345–364). Hillsdale, NJ: Erlbaum.

Schreuder, R., & Baayen, H. (1997). How complex simplex words can be. *Journal of Memory and Language*, 37, 118–139. https://doi.org/10.1006/jmla.1997.2510

Schreuder, R., Burani, C., & Baayen, H. (2003). Parsing and semantic opacity. In E. Assink, & D. Sandra (Eds.), *Reading complex words. Cross–language studies* (pp. 159–189). Dordrecht: Kluwer.

Selkirk, E. (1982). *The syntax of words*. Cambridge, MA: The MIT Press.

Sims, A., & Parker, J. (2015). Lexical processing and affix ordering: cross–linguistic predictions. *Morphology*, 25, 143–182. https://doi.org/10.1007/s11525-015-9257-0

Szymanek, B. (2015). Remarks on tautology in word–formation. In L. Bauer, L. Körtvélyessy, & P. Štekauer (Eds.), *Semantics of complex words* (pp. 143–161). Springer.

Taft, M. (1979). Lexical access via an orthographic code: the basic orthographic syllabic structure (BOSS). *Journal of Verbal Learning and Verbal Behavior*, 18, 21–39. https://doi.org/10.1016/S0022-5371(79)90544-9

Taft, M. (1981). Prefix striping revisited. *Journal of Verbal Learning and Verbal Behavior*, 20, 289–297. https://doi.org/10.1016/S0022-5371(81)90439-4

Taft, M. (2003). Morphological representation as a correlation between form and meaning. In E. Assink, & D. Sandra (Eds.), *Reading complex words: cross–linguistic studies* (pp. 113–137). Amsterdam: Kluwer.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effects. *The Quarterly Journal of Experimental Psychology*, 57A, 745–765. https://doi.org/10.1080/02724980343000477

Taft, M., & Ardasinski, S. (2006). Obligatory decomposition in reading prefixed words. *The Mental Lexicon* 1(2), 183–199. https://doi.org/10.1075/ml.1.2.02taf

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, 15, 607–620. https://doi.org/10.1016/0022-5371(76)90054-2

Taft, M., & Kougious, P. (2004). The processing of morpheme–like units in monomorphemic words. *Brain and Language*, 90, 9–16. https://doi.org/10.1016/S0093-934X(03)00415-2

Talamo, L. (2015). Suffix combinations in Italian: Selectional restrictions and processing constraints. In S. Manova (Ed.), *Affix ordering across languages and frameworks* (pp. 175–204). Oxford: Oxford University Press.

Valera, S. (2015). Conversion. In P. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), *Word–formation. An international handbook of the languages of Europe*. Vol. 1 (pp. 322–339). Berlin: Mouton de Gruyter.

Zirkel, L. (2010). Prefix combinations in English. *Morphology*, 20(1), 239–266. https:/doi.org/10.1007/s11525-010-9151-8

Ograničenja pri proširivanju afiksalnih kombinacija: strukturna ograničenja i uvjeti procesiranja

Istraživanja o mentalnom leksikonu podržana su analizama o tome na koji su način složene riječi mentalno reprezentirane i procesirane. U ovome radu analizira se proširivanje višestruke afiksacije na sintagmatskoj razini; preciznije, analizira se procesiranje složenih riječi koje sadržavaju četiri sufiksa koja nalazimo u obrascima tvorbe riječi u portugalskom jeziku. Premda je pojedinačno dodavanje sufiksa u skladu sa strukturnim ograničenjima, kombinacija više sufiksa rezultira složenim riječima niske čestotnosti i niskog očekivanja kod govornika, što pridonosi smanjenoj semantičkoj transparentnosti i istaknutosti afikasa u kombinaciji. Ova analiza pokazuje povezanost izmedu navedenih čimbenika i iskustva govornika s pojedinom afiksalnom kombinacijom, što određuje činjenica da su kombinacije građene na principu obrazaca. U radu predlažemo da jedan sufiks priziva ostale sufikse dokle god je kombinacija sufikasa očekivana. Rjeđe heterokategorijske složene riječi s kombinacijom četiriju sufiksa kontrastiraju se s rjedim riječima koje sadržavaju pleonastičku afiksaciju. U potonjem tipu riječi redundantnost značenjskih struktura povećava značenjsku transparentnost riječi, što upućuje na predvidivost koja se odvija na značenjskoj razini afiksalne kombinacije. Procesiranje složenih riječi ovisi o razini govornikova očekivanja prema kombinaciji afikasa, što ograničava razinu prihvaćanja riječi kod govornika.

Keywords: affixal combination, word-formation patterns, semantic transparency, mental lexicon, Portuguese

Ključne riječi: afiksalna kombinacija, obrasci tvorbe riječi, značenjska transparentnost, mentalni leksikon, portugalski

Appendix A

To understand the pertinence of displaying Google results, consider the caveat provided in section 5.1.2, where we have explained that CRPC frequencies do not reflect the usage of slang, oral and informal words, such as diminutives.

Neutral words	CRPC Frequency per million of words	Number of Results in Google
viável	11.05	5 670 000
paralelo	15.83	36 600 000
igualmente	129.64	66 100 000
rápido	30.09	339 000 000
comparação	24.41	36 800 000
bué	0.85	2 580 000
pão	30.92	59 900 000
cor	35.92	554 000 000
paciência	15.7	34 200 000
construção	231	181 000 000
senhor	118.5	108 000 000
espacial	9.65	43 500 000
amarelo	12.56	124 000 000
livro	104.9	213 000 000
porra	0.63	40 900 000
desejável	18.34	7 100 000
referir	99.7	9 990 000
céu	37.64	120 000 000
nicles	0.07	109 000
camisa	6.77	175 000 000
imediatamente	56.17	36 900 000
feliz	34.34	601 000 000
cavaleiro	17.15	16 400 000
coração	52.35	240 000 000
ler	92.16	223 000 000
utilidade	33.45	33 200 000
espelho	11.5	77 200 000

Pseudo–words
quirilete
mitoloco
vontoco
citisco
agulhisto
socir
ricombar
arcadilo
pomponipom
fanti
cantaco
livivi

Evaluative non–frequent words	CRPC Frequency per million of words	Number of results in Google
cadelinhazinha	0	0
gatinhazinha	0	757
pequeninozinho	0	160
copinhozinho	0	29
cãozinhinho	0	7
casinhotozinho	0	1
lapinhozito	0	2
menininhazinha	0	399
mesinhitazinha	0	0
pequenininhozinho	0	193
cavalinhozinho	0	10
riachozinho	1 match in 1 text 0 frequency	811
aranhiçozinho	0	3

Evaluative frequent words	CRPC Frequency per million of words	Number of results in Google
passarinho	1.1	38 600 000
cafezinho	0.26	8 630 000
pequenino	2.65	5 900 000
pequenito	0.12	4 120 000
amarelinho	0.05	3 630 000
livrinho	0.97	3 210 000
carrinho	1.31	90 500 000
casinha	1.08	53 700 000
caderninho	0.14	4 930 000
maçãzinha	0	149 300
bichinho	0.57	17 500 000
florinha	0.04	103 000
escadinha	0.12	2 310 000

(–al–/bil–)–iz(a)–bil–idade words	CRPC Frequency per million of words	Number of results in Google
globalizabilidade	0	0
estabilizabilidade	0	2
contratualizabilidade	0	0
comercializabilidade	0	26
legalizabilidade	0	156
familiarizabilidade	0	0
materializabilidade	0	0
socializabilidade	0	2
contabilizabilidade	0	374
compartibilizabilidade	0	3

Appendix B presents the sentences used as stimuli in the Recall task.

Neutral words	Sentence	Source
viável 'viable'	*Considero viável um diálogo entre ambas as disciplinas.* 'I consider viable a dialogue between both subjects.'	*https://books.google.pt/books?id=uM2DbHFotKAC&pg=PA22&lpg=PA22&dq=considero+vi%C3%A1vel&source=bl&ots=pgrY41GjRa&sig=lwRt9bkKFDYuqaUPtH4X–3udt7A&hl=pt–PT&sa=X&ved=0ahUKEwic2LW568HTAhWLWhoKHeMbDxEQ6AEIODAE#v=onepage&q=considero%20vi%C3%A1vel&f=false*
paralelo 'parallel'	*Este sítio não tem paralelo em Portugal.* 'This place is unparalleled in Portugal.'	Linguateca: par=ext11354–clt–94b–1
igualmente 'likewise'	*Esta verba será igualmente atribuída aos familiares.* 'This sum will be likewise attributed to the relatives.'	Linguateca: par=ext307084–soc–93a–2
rápido 'fast'	*Oferece grandes possibilidades de um rápido desenvolvimento.* '[It/She/He] offers great possibilities of a fast development.'	CRPC: A13096
comparação 'comparison'	*Machete considera «absurda» comparação entre Portugal e Grécia.* 'Machete considers «absurd» the comparison between Portugal and Greece.'	http://www.aeiou.pt/quiosque/machete–considera–absurda–comparacao–das–situacoes–portuguesa–e–grega
bué 'very'	*Ele anda bué de chateado.* 'He is very annoyed.'	Linguateca: par=ext1116514–clt–94b–2
pão 'bread'	*No saco havia nacos de pão [...] e uma garrafa de água.* 'In the bag there were pieces of bread and one bottle of water.'	Linguateca: par=ext248–nd–91a–1
cor 'colour'	*Até os vendedores mudaram a cor às pipocas.* 'Even the sellers have changed the colour of popcorn.'	Linguateca: par=ext1271072–pol–95b–2
paciência 'patience'	*Peço paciência para o tom subjectivo deste texto.* 'I ask your patience for the subjective tone of this text.'	Linguateca: par=ext1046–soc–93a–2

construção 'construction'	*Não haverá o actual processo de construção europeia.* 'the current process of European construction will be amiss.'	Linguateca: par=ext558614–nd–91a–2
senhor 'mister'	*O senhor Gabriel [...] é bem conhecido pelo público português.* 'Mister Gabriel is well known by the Portuguese public.'	Linguateca: par=ext1335836–soc–98a–2
espacial 'spatial'	*O programa espacial japonês vive tempos de crise e de críticas.* 'The Japanese spatial programme is undergoing times of crisis and critics.'	Linguateca: par=ext4243–clt–98a–3
amarelo 'yellow'	*O palco [...] tem um gigantesco arco amarelo de 100 pés.* 'The stage [...] has a gigantic yellow 100 feet arch.'	Linguateca: par=ext45180–clt–97a–2
livro 'book'	*A Amnistia Internacional considera–o um livro sério de depoimentos.* 'Amnesty International considers it a serious book of statements.'	Linguateca: par=ext44940–pol–95b–1
porra (literally: 'cum') 'bleeding'	*Não havia porra nenhuma de transparência naquele país.* 'There was no bleeding transparency in that country.'	Linguateca: par=ext326344–nd–91a–2
desejável 'desirable'	*Seria desejável uma maior problematização de conceitos.* 'A greater problematisation of concepts would be desirable.'	Linguateca: par=ext6362–nd–91a–2
referir 'to refer'	*Mas está a referir–se a isto.* 'But [he/she] is referring to this.'	CRPC: noCOD_1016299
céu 'sky'	*E observou–se o movimento do céu, do mar, das cagarras.* 'And we observed the movement of the sky, of the sea, of the Cory's shearwaters.'	Linguateca: par=ext4971–nd–91b–1
nicles 'zilch'	*Não percebe nicles de comédia.* '[He/she] understands zilch of comedy.'	https://agranelemaisbarato.blogspot.pt/2007/09/repararam–na–constituio–das–equipas.html
camisa 'shirt'	*Noutro dia, trouxeram–me uma camisa da Tailândia.* 'Once, they brought me a shirt from Thailand.'	Linguateca: par=ext29229–eco–98b–1

imediatamente 'immediately'	*São imagens que imediatamente nos surgem.* '[Those] are images that immediately come to our mind.'	Linguateca: par=ext110650–clt–95a–1
feliz 'happy'	*O empate traduz de forma feliz um jogo que ficou no meio.* 'The tie shows in a fortunate way a game that ended halfway through'.	Linguateca: par=ext18–des–94a–2
cavaleiro 'knight'	*O próprio cavaleiro acreditava nisso.* 'The knight himself believed it.'	Linguateca: par=ext81836–des–92b–2:
coração 'heart'	*Tem uma alma de aço e um coração apaixonado.* '[He/She] has a soul of steel and a passionate heart.'	Linguateca: par=ext472005–soc–93b–1
ler 'to read'	*«A Imperatriz», que acabo de ler, é apaixonante.* '«The Empress», which I have just read, is enthralling.'	Linguateca: par=ext7925–opi–97a–2
utilidade 'usefulness'	*Qual é então a utilidade da portaria?* 'What is the usefulness of the decree?'	Linguateca: par=ext11291–clt–soc–93a–1
espelho 'mirror'	*Ontem comprei um espelho do século dezoito.* 'Yesterday I bought an 18th century mirror.'	https://books.google.pt/books?isbn=1310024782

Pseudo–words	Sentences	source
quirilete	*A Joana comprou um quirilete magnífico.* 'Joana bought a magnificent quirilete.'	author
mitolo	*Do mitolo resta apenas um pedaço.* 'Of the mitolo only one piece is left.'	author
vontoco	*Pensa no vontoco com carinho.* 'Think about he vontoco with tenderness.'	author
citisco	*Há um citisco dentro do armário.* 'There is a citisco inside the wardrobe.'	author
agulhisto	*O Rui é demasiado agulhisto nas suas decisões.* 'Rui is too agulhisto in his decisions.'	author
socir	*Continuou a socir durante a manhã.* '[She/He/It] still socir during the morning.'	author
ricombar	*A partir do ricombar avistam–se as planícies.* 'From the ricombar the plains are sighted.'	author
arcadilo	*Observou todo o arcadilo velho.* '[She/He] observed the entire old arcadilo.'	author
pomponipom	*Trouxe o pomponipom durante todo o dia.* '[She/He] wore the pomponipom all day long.'	author
fanti	*Considera esta fanti detestável.* '[She/He] considers this fanti detestable.'	author
cantaco	*Apreciou o cantaco sem interesse particular.* '[She/He'] appraised the cantaco without particular interest.'	author
livivi	*Um livivi pode viver 100 anos.* 'A livivi may live 100 years.'	author

Evaluative non–frequent words	sentence	Source
cadelinhazinha 'little, little/cute, cute bitch'	*O João tem uma cadelinhazinha muito gira.* 'João has a little, little/cute, cute very cute she–dog.'	author
gatinhazinha 'little, little/cute, cute cat'	*Uma gatinhazinha [...] miava com toda a força.* 'A little, little/cute, cute cat [...] meowed resolutely.'	http://insharee.com/u/36534167?max_id=1217671302979068136_36534167
pequeninozinho 'little, little one'	*Sim, meu pequeninozinho, faz queixinhas à vó.* 'Yes, my little, little one, complain to grandma.'	http://oreinabarriga.blogspot.pt/2005/11/?m=0
copinhozinho 'small, small glass'	*Aquele copinhozinho de café me custou uns 6 goles.* 'That small, small glass of coffee costed me 6 sips.'	http://donaredonda2012.blogspot.pt/2014/03/dona–redonda.html
cãozinhinho 'little, little /cute, cute dog'	*Sou um cãozinhinho muito fofo.* 'I am a very cute little, little/cute, cute dog.'	https://www.facebook.com/maxantoniospitz/
casinhotozinho 'very small house'	*Vivia num casinhotozinho perto do mar.* 'She/He lived in a very small house by the sea.'	Author. The word is listed in www.infopedia.pt
lapinhozito 'little, little pencil'	*Roeu o lapinhozito até ao fim.* 'She/He gnawed at the little, little pencil till the last.'	https://www.facebook.com/mariflorsa/
menininhazinha 'little, little/cute, cute girl	*Fotos de família [...] mostram [...] uma menininhazinha com laços na cabeça.* 'Photos of the family [...] show a little, little/cute, cute girl with ribbons in the head.'	http://gcn.net.br/mobile/noticia/317092/opiniao/2016/04/lembrancas
mesinhitazinha 'little, little, little table'	*Comprou uma mesinhitazinha de madeira.* 'She/He bought a little, little, little wooden table.'	author
pequenininhozinho 'very, very, very small'	*Ele é muito pequenininhozinho, portátil, confiável.* 'He is very, very, very small, portable, trustable.'	http://sentimentosreciclados.blogspot.pt/2010/04/cavalinhozinho.html

cavalinhozinho 'little, little/cute, cute horse'	*Encontrei esse cavalinhozinho branco na calçada.* 'I found that little, little/cute, cute horse by the sidewalk.'	http://www.imgrum.org/user/tatipirilim/1329250672/1376497795787959248_1329250672
riachozinho 'small rivulet'	*Era um riachozinho onde as pessoas se banhavam no Verão.* 'It was a small rivulet where people bathed in the summer.'	CRPC: J38354
aranhiçozinho 'very small spider"	*Lá andava o aranhiçozinho a rolar.*[7] 'There was the very small spider rolling on.'	http://www.patrolaventura.com/forum/viewtopic.php?p=405517

Evaluative frequent words	sentences	source
passarinho 'small bird'	*Contam a história do passarinho que tinha um segredo.* 'They tell the story of the small bird who had a secret.'	Linguateca: par=ext222379–clt–soc–92a–2
cafezinho 'diminutive of coffee'	*Fui tomar um cafezinho em Campinas.* 'I went for a coffee in Campinas.'	https://gostososemgluten.com.br/2016/05/24/tomando–um–cafezinho–descobri–um–sufle–de–chocolate–sem–gluten–delicioso/
pequenino 'very small'	*Então vamos lá fazer um pequenino teste.* 'So let us make a very small test.'	Linguateca: par=ext24190–nd–92a–2
pequenito 'very small'	*Mas o pequenito respirava com custo.* 'But the little one breathed with difficulty.'	Linguateca: par=ext617782–soc–92a–1
amarelinho 'very yellow'	*O arroz–doce estava [...] amarelinho e com um pouco de canela.* 'The rice pudding was very yellow and with a bit of cinnamon.'	CRPC: J67188
livrinho 'little book'	*Seguiram à risca o livrinho de instruções.* 'They followed the little instruction book to the letter.'	Linguateca: par=ext13456–des–92a–2

7 We have changed the original sentence to get a simpler syntactic structure. The original sentence was: *É sempre bom ver esse aranhiçozinho a rolar.* 'It is always good to see that very small spider rolling on.'

carrinho 'nice car'	*O seu belo carrinho sofreu umas amolgadelas.* 'Her/His beautiful nice car sustained some dents.'	Linguateca: par=ext61553–pol–94a–3
casinha 'small/nice house'	*Arranjem uma casinha para a gente.* 'Get us a small/nice house.'	Linguateca: par=ext332977–soc–94b–1
caderninho 'small notebook'	*Ando sempre com um caderninho onde aponto pequenas frases.* 'I always bring a small notebook with me where I note down some small sentences.'	Linguateca: par=ext103761–clt–94b–1
maçãzinha 'small/nice apple'	*Comi uma maçãzinha que estava deliciosa.* 'I ate a small/nice apple that was delicious.'	oprimeiroblogdomartim.blogspot.fr/2009/10/minha–1–frutinha.html
bichinho 'small animal'	*A joaninha não é só um bichinho colorido e simpático.* 'The ladybird is not just a colourful and nice small animal.'	Linguateca: par=ext9835–nd–91b–4
florinha 'small flower'	*Há uma florinha branca que descobri na minha infância.* 'There is a white small flower that I discovered during my childhood.'	olhaioliriodocampo.blogspot.com/2009/03/florinha–branca.html
escadinha 'little stairs'	*Iam subindo a escadinha de ferro.* 'They were climbing up the little iron stairs.'	CRPC: L0379

(–al–/bil–)–iz(a)–bil–idade words	Sentences	Source
globalizabilidade 'quality of being able to be globalised'	*A globalizabilidade das notícias é hoje muito rápida.* 'The 'quality of being able to be globalised' of the news is very fast today.'	author
estabilizabilidade 'quality of being able to be stabilised; stabilisability'	*Resgatam os conceitos de estabilizabilidade [...] do caso clássico.* '[They] recovered the concepts of stabilisability of the classic case.'	https://pdfs.semanticscholar.org/e84e/906adf3407be866121ba486e06bb7d5442c7.pdf
contratualizabilidade 'rate of being able to be contracted'	*A contratualizabilidade tem decrescido anualmente.* 'The 'rate of being able to be contracted' has decreased annually.'	author
comercializabilidade 'tradability'[8]	*O grau de "comercializabilidade" [...] é insuficiente.* 'The degree of tradability is insufficient.'	bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/1040/2202.pdf
legalizabilidade 'quality of being able to be legalised'	*A intervenção na questão de legalizabilidade significa garantir condições.* 'The intervention in the issue of the 'quality of being able to be legalised' means to guarantee conditions.'	http://www.portaldoconhecimento.gov.cv/bitstream/10961/3971/1/DISSERTA%C3%87%C3%83O_Mest_Kathya_10_011%201%20(2).pdf
familiarizabilidade 'quality of being able to get familiar'	*A familiarizabilidade do João com a Ana vai demorar tempo.* 'The 'quality of' João 'being able to get familiar' with Ana will take much time.'	author
materializabilidade 'quality of being able to be materialised'	*Um sonho tem uma materializabilidade dependente de muitos fatores.* 'A dream has a 'quality of being able to be materialised' dependent upon many factors.'	author

8 Translation presented at the source (bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/1040/2202.pdf): «Em muitos setores o grau de "comercializabilidade" (tradability) é insuficiente.».

socializabilidade quality of being able to be socialised'	*Doam os filhotes sem nenhuma socializabilidade ou imunidade.*[9] 'They offer the puppies with neither 'quality of being able to be socialised' nor immunity.'	patriciapiedadesouza.blogspot.com/2010_04_01_archive.html
contabilizabilidade 'quality of being able to be counted'	*Pesquisas não registradas [...] não apresentam contabilizabilidade [nenhuma].*[10] 'Non–registered searches do not present any 'quality of being able to be counted'.'	http://www.avozdocampo.com/cidades/araci/para–ter–coabilidade–pesquisas–eleitorais–precisam–de–registro–na–justica–e–sao–de–acesso–publico/
compartibilizabilidade 'quality of being able to be divided'	*As construções recentes mostram uma compartibilizabilidade considerável.* 'Recent constructions show a considerable 'quality of being able to be divided'.'	http://lusios.blogspot.pt/

9 The original sentence is: *Doam os filhotes sem nenhuma imunidade ou socializabilidade.* We have changed the position of imuni*dade w*ith *socializabilidade*, so that the target word was not the last word of the sentence.

10 The original sentence is: *Especialistas afirmam que pesquisas não registradas e, portanto, não autorizadas pela justiça eleitoral não apresentam contabilizabilidade.* Apart from deleting some elements, we have added the word *nenhuma* 'none', so that the target word was not the last word of the sentence.